DOCUMENT RESUME

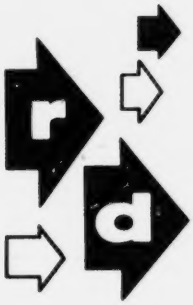
ED 077 875 SP 006 556

TITLE	Marking and Reporting Pupil Progress. Research Summary 1970-71.
INSTITUTION	National Education Association, Washington, D.C. Research Div.
PUB DATE	70
NOTE	55p.
AVAILABLE FROM	National Education Association, 1201 Sixteenth Street, N.W., Washington, D.C. 20036 (Stock No. 434-22854, $1.25)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	Academic Achievement; *Academic Records; Achievement Rating; Grades (Scholastic); *Grading; Rating Scales; *Report Cards; *Student Development; Student Evaluation; *Student Records

ABSTRACT

This is a research summary of aspects of marking and reporting pupil progress: practices, philosophies of marking, purposes, assets, and liabilities of various marking and reporting practices. The summary is divided into three sections: a) determining and reporting pupil progress (included are descriptions of marking purposes and practices and factors that influence marking), b) methods of reporting, and c) research on marking and reporting. Listed in this last section are statistics and descriptions of various studies; the latter cover the grading of arithmetic, the grading behavior of a college faculty, grades as they reinforce contingencies and attitude changes, student reactions to course grades, and many others. The summary concludes with a plea for a greater understanding of the reportage system and the need for more attention given to the content (rather than the method) of reports. (JA)

ED 077875



RESEARCH SUMMARY 1970-S1

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICRO FICHE ONLY HAS BEEN GRANTED BY

NEA

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Marking and Reporting Pupil Progress

SP 000 538

RESEARCH DIVISION – NATIONAL EDUCATION ASSOCIATION

Copyright© 1970 by the
National Education Association
All Rights Reserved

FILMED FROM BEST AVAILABLE COPY

NATIONAL EDUCATION ASSOCIATION

GEORGE D. FISCHER, President
SAM M. LAMBERT, Executive Secretary
GLEN ROBINSON, Assistant Executive Secretary for Research

RESEARCH DIVISION

GLEN ROBINSON, Director

SIMEON P. TAYLOR III, Assistant Director
WILLIAM S. GRAYBEAL, Assistant Director
ALTON B. SHERIDAN, Assistant Director
FRIEDA S. SHAPIRO, Assistant Director
JEAN M. FLANIGAN, Assistant Director

GERTRUDE N. STIEBER, Senior Professional Associate

DONALD P. WALKER, Professional Associate

RICHARD E. SCOTT, Chief Statistician

VALDEANE RICE, Administrative Associate

ELIZABETH C. MOFFATT, Senior Staff Associate
MARSHA A. REAM, Senior Staff Associate
JEANETTE G. VAUGHAN, Senior Staff Associate

ANN T. McLAREN, Staff Associate
SANDY M. PEMBERTON, Staff Associate
KENNETH L. SANDVIG, Staff Associate
NINA C. SIMMONS, Staff Associate

GRACE BRUBAKER, Chief, Information
FRANCES H. REYNOLDS, Librarian
WALLY ANNE SLITER, Chief, Copy Preparation

HELEN KOLODZIEY, Assistant Chief, Information
MOLLY B. TEMPLETON, Assistant Librarian

BEATRICE C. LEE, Publications Editor

Research Summary 1970-S1: MARKING AND REPORTING PUPIL PROGRESS

Project Director: MARSHA A. REAM, Senior Staff Associate

Subscription Rate: One-year subscription to NEA Research Division Reports and Summaries, $18; send inquiries to NEA Records Division.

Reproduction: No part of this Summary may be reproduced in any form without written permission from the NEA Research Division, except by NEA Departments and affiliated associations. In all cases, reproduction of the Research Summary materials must include the usual credit line and the copyright notice. Address communications to the Publications Editor.

Price of Summary: Single copy, $1.25. Stock #434-22854. Discounts on quantity orders: 2-9 copies, 10%; 10 or more copies, 20%. Orders not prepaid will be billed with shipping charges added. Order from Publications Sales Section and make checks payable to the National Education Association, 1201 Sixteenth Street, N. W., Washington, D. C. 20036.

CONTENTS

Foreword 4

Highlights 5

Determining and Reporting Pupil Progress 7

- Purposes of Marking 8
- Bases for Marking and Reporting 12
- Evidence of Pupil Progress 14
- Determining the Meaning of a Mark 15
- Factors That Influence Marking 16
- Summary 17

Methods of Reporting 19

- Letter-Number Systems 19
- Pass-Fail Systems 22
- Check Lists and Rating Scales 22
- Anecdotal Records 23
- Behavior Description Method 23
- Correspondence with Parents 24
- Teacher-Parent Conferences 24
- Self-Evaluation 26
- Summary 26

Research on Marking and Reporting 28

- Current Practice 31
- Recent Research, 1960 to the Present 32
- Studies of Marking and Reporting in the Elementary and Secondary School 33
- Studies of Marking and Reporting at College Level 37
- Summary 45

Selected References 47

FOREWORD

MARKING AND REPORTING PUPIL PROGRESS are perennial topics of discussion among teachers, parents, and pupils; everyone has an opinion. Some have even suggested abolishing school marks. Although teachers would probably rank marking and reporting at the bottom of their list of preferred activities, there may be no other activity that has greater potential for interpreting the school program, for securing cooperation between home and school, and for promoting pupil development. Despite many years of experimentation and research, the problems and issues related to marking and reporting remain unsettled. To assist interested persons in understanding the multiplicity of factors related to marking and reporting, the NEA Research Division undertook the task of assembling and summarizing the results of research and opinion.

This research summary discusses reporting practices, the philosophies of marking, the purposes of marking, and the assets and liabilities of various marking and reporting practices. Emphasis, however, is placed on the findings of representative research studies.

It is hoped that this report will be a useful source book to direct attention to representative inquiries and to stimulate thinking concerning marking and reporting pupil progress.

GLEN ROBINSON
Director, Research Division

HIGHLIGHTS

NO UNIFORM METHOD of marking pupils and reporting their progress to parents exists among school systems in the United States. The most common technique of reporting is the report card, a term which is gradually being supplanted by progress report. Even among report cards there is variation from one school system to another.

A review of the research and experimentation with regard to grading and reporting suggests the cyclical nature of the various grading and reporting practices. Some older practices, e.g., anecdotal records, appear to be enjoying a renascence of sorts. A basic problem in the search for a viable grading and reporting system is meeting the need for both descriptive and quantitative information regarding pupil progress.

Research and experimentation have indicated that the following appear to be some general changes in current practice.

Practice:	Some change toward:
1. Pupil evaluation conducted and reported exclusively by the teacher.	Pupil preparation of a brief written evaluation of their own work.
2. Listing only broad subject fields. The pupil receives a single mark in each broad subject classification.	Explaining in terms of pupil behavior the activities that compose each subject-matter field as well as character traits. More detailed explanation of the meaning of marks on the written report card.
3. Comparing all pupils with a set standard or with their classmates.	Combined use of individual progress marks and comparative marks. Comparing each pupil's progress with his own apparent ability (especially in elementary school) or with himself as well as with others.
4. Using a single report form for the entire school.	Developing forms suited specifically to the goals of specific grades or levels.
5. Pupils receiving marks at six-week or nine-week intervals.	Less frequent reporting, e.g., twice a year.
6. Providing a line for a teacher comment, and a line for the parent's signature.	Providing space for both teacher and parent comments.
7. Central office supervisory staff, with suggestions from teachers, develop the report card.	Organizing committees of supervisors, teachers, parents, and pupils to improve reporting practices.
8. Use of only a card to report pupil progress to parents.	Combining reporting by both report cards and parent-teacher conferences, often alternating for reporting periods.
9. Using a relatively small card printed black on white.	Using a large folder in color with typographical design and explanations to parents. Pictures and cartoons on the card are often used at the primary level.

10.	Including only school subjects and perhaps an item titled character or deportment.	Including numerous objectives under such titles as social adjustment, personal development, and work habits.
11.	Using percentage or letter grades, which are sometimes defined in such terms as excellent, good, fair, and failure.	Developing additional symbols or statements, which reflect more understanding of child development (such as needs more time and help). Increased use of descriptive terms and interpretation.

Rather than use a report card, a few schools prefer the informal letter to parents or the teacher-parent conference for reporting pupil progress. These two techniques can provide individualized descriptions of a child's behavior but also demand considerably more teacher time.

Many plans for marking and reporting are combinations of various methods geared to the need of the local school or school system. Many local schools or school systems are seeking to involve the parents as active participants in the pupil evaluation process. Cooperative participation is considered a key concept in efforts to improve marking and reporting. Educators, parents, and pupils usually study the school's educational objectives and then determine the most effective way to communicate the results of pupil progress in terms of them. Only after careful study of various approaches, past experiments, critical reactions, and revisions can a school hope to effect a defensible system of marking and reporting.

DETERMINING AND REPORTING PUPIL PROGRESS

ONE OF THE GREATEST concerns of teachers, parents, and pupils is determining and reporting pupil progress. The fundamental purpose of any method of determining and reporting pupil progress is to establish a flow of communication between the school, the pupil, and the home in the interest of providing feedback and guidance so that each pupil may achieve to the degree that his talent will allow. To accomplish its fundamental task, evaluation must perform the following: (a) facilitate self-evaluation; (b) encompass all the objectives; (c) facilitate teaching and learning; (d) generate records appropriate to various uses; and (e) facilitate decision-making on curriculum and educational policy (6:234).[1/] Whether or not this fundamental purpose has been fully achieved remains a moot question.

Marking and reporting have progressed since the earliest report cards in American education. These early reports contained simple little messages and were generally quite colorful and artistic. Comments about achievement in subjects such as penmanship, spelling, and reading developed later. These evaluations, together with a report on attendance and deportment, persisted for many years. By the turn of the century, however, the two most widely used systems were the percentage grade and the five-point scale, the latter often expanded to 15 points by the use of pluses and minuses.

With the growth of the standardized testing movement at the beginning of the twentieth century, grades came in for much criticism on the grounds that they were unreliable measures. Early studies by Starch, Elliott, and others in 1912 indicated the low reliability of both letter and percentage grades (99:391). Between about 1918 and 1928, much emphasis was placed on the development of reliable measurement devices to ensure accuracy and on the standardization of marking systems. A study by Odell in 1925 revealed that among 281 Illinois schools, almost 100 varieties of marking systems prevailed (111:499). A 1928 proposal (111:500) to render marks more reliable suggested that they represent achievement only; that they be based solely on the results of reliable, objective tests as far as possible; that conscientious effort, daily work, and recitation not be used in determining them; and that the normal curve be used as a guideline for the distribution of marks. The author believed that the necessary pedagogical, administrative, and guidance functions of marks were not really met by the then current systems of letters and percentages, and that the remedy for this defect was not to abandon the marks, but to make them reliable measures (111:499). Percentage grades were considered especially undesirable because, according to Odell, research had shown that teachers could not make such fine distinction as percentage grades required nor could there be rigid definition of the three reference points--0, 100, and the passing mark (12:26). Moreover, percentage grades conveyed, because of their numerical form, a false impression of accuracy (81:458).

As the progressive movement in education began in the late 1920's, the concerns about grading shifted ground. The emphasis changed from a "let's make grades more reliable" approach to a concern that marks, reliable or unreliable, were a poor evaluative device because a single grade conveyed so little information about strengths and weaknesses in a given area. Moreover, a number of studies indicated that marks fostered unhealthy competition among students and often had an undesirable effect upon the mental health of children (12:25), frequently producing a level of anxiety that seriously handicapped the child in his work (73:29).

With these new concerns began a new approach to grading, the evaluative approach. In 1933, the Newton, Massachusetts, public schools abolished grade cards and replaced them with a private note to parents, discussing the child (92:225). The Eight Year Study of the Progressive Education Association provided the opportunity for a full-scale program of evaluation, an attempt to put an evaluation rationale into practice. It showed that testing specialists had been concerned too long with the knowledge aspects of education and had not placed enough emphasis on the intangible outcomes of the educative process (99:6). The comprehensive report on evaluation which grew out of the Eight Year Study rested upon three basic conditions: (a) that evaluation and recording must be directly related to each school's purposes and philosophy; (b) that an

1/ The numbers in parentheses refer to items in the bibliography.

evaluation program must be comprehensive, including appraisal of progress toward all the major objectives of the school; and (c) that teachers should participate in the construction of all instruments for evaluation and forms for records and reports (102:18).

The 10-year study at the Campus Research-Laboratory Schools of Colorado State College of Education is described in Wrinkle's Improving Marking and Reporting Practices (121) which has become a classic on the subject. Wrinkle believed that the use of the single letter marking system is supported by six fallacies: (a) that the mark conveys information effectively, (b) that anyone can achieve any mark if he makes the necessary effort, (c) that people succeed outside the school about the same as they do in school, (d) that the mark is comparable to a pay check, (e) that marking provides a justifiable introduction to competitive adult life, and (f) that the mark can be used as a means without eventually being considered as an end in itself (121:49).

The evaluation approach to marking and reporting dominates the current literature in education. The 1967 Yearbook of the Association for Supervision and Curriculum Development calls for a complete revision of the going systems of evaluation, which, the Yearbook Committee believes, have drifted largely into the service of marking, grading, and crediting. Such a revision would require a probing examination of curriculum objectives and would result in a more diagnostic approach to teaching (6).

Today, there is a definite trend toward revising methods of marking and reporting. Child psychology and a redefining of the school's objectives have had their influence. Attention has been centered on the whole child. His achievement in relation to his own potential for development is emphasized. The complexity of this kind of evaluation is reflected in changed methods of reporting. Such devices as dual marking, personal letters, and teacher-parent conferences are now being used in many instances.

Purposes of Marking

In his book, Wrinkle discussed the functions of marks and listed four functions that marks are intended to fulfill, and although other statements have appeared over the years since then, no essential alteration in this listing has occurred.

1. Administrative functions: Marks indicate whether a student has passed or failed, whether he should be promoted or required to repeat the grade or course, and whether he should be graduated. They are used in transferring a student from one school to another and in judging candidates for admission to college. They may be used by employers in evaluating prospective employees.

2. Guidance functions: Marks are used in guidance and counseling in identifying areas of special ability and inability, in deciding on the advisability of enrolling the student in certain courses and keeping him out of others, and in determining the number of courses in which he may be enrolled.

3. Information functions: Marks are the chief means employed by the school in giving information to students and their parents regarding the student's achievement, progress, and success or failure in his schoolwork.

4. Motivation and discipline functions: Marks are used to stimulate students to make greater effort in their learning activities. They are used for the same purpose in determining eligibility to honors of many different kinds such as participation in school activities, eligibility to play on the team, membership in selected groups, the winning of scholarships, etc. (121:31-32).

These four classifications are not mutually exclusive; they overlap. For instance, the use of marks for awarding a scholarship provides motivation, but it also serves an administrative function. Thus, almost any one function may be under more than one classification (121:32).

Motivation

One purpose of reports is to encourage pupils to greater effort in school. It is argued that fear of failure produces greater achievement. If the pupil knows that passing from grade to grade is automatic, he will not work. Research, however, does not consistently support these lines of reasoning. For instance, one study compared nine elementary school systems employing "social promotion" and nine elementary school systems maintaining rigid standards of promotion. No significant differences were found either in achievement or in intelligence between seventh-grade groups. The range of abilities aggravated the instruction program in the schools where failure was common. In fact, grade standards tended to be lowered (1:547).

Conversely, it has been found that certain pupils will be challenged to work for higher marks. Interest in a particular subject may develop through feelings of accomplishment and success. There are, however, some detriments to motivation in working for marks per se. Creativity and imagination may be curbed when

a pupil conforms rigidly to a set of standards for a mark. A pupil may lack the mental ability to achieve the set standard and is not motivated to learn when despite his efforts he cannot succeed. On the other hand, if marks represent improvement rather than status alone, the pupil may have an incentive for which to strive.

Students react differently to marks. For instance, a child from a home where school success is not considered important may pay little or no attention to the prospect of a low grade. For the child whose parents consider school success vital to "getting ahead in the world," the prospect of a low mark and consequent parental disapproval may cause the pupil to exert greater effort to do well.

The way a student's motivation is affected by grading systems can vary according to (a) the standard of comparison, (b) parents' and friends' attitudes toward marks, and (c) the teacher's emphasis on marks. A brief inspection of each of these variables follows.

1. The standard of comparison--A mark is a judgment of a student's progress compared with some standard. There are three principal kinds of standards. First, a student may be compared with his classmates. Second, he may be compared with a teacher-set standard. As a third possibility, the pupil's present success can be compared with his own past performance, regardless of the level of work being done by his classmates.

The first two of these standards have been the most popular. The third is gaining adherents, especially in the elementary grades. Sometimes a teacher uses more than one of these approaches. Pupil motivation may be affected by the particular standard the teacher chooses to use.

It is probable that the somewhat above-average pupil is stimulated to work well when he is being rated against the others' performances. However, under the system of comparing the pupil with his classmates, it is possible for the capable student in class to stay at or near the top without exerting much effort. If the capable student can easily outstrip his classmates, he is not likely to work up to his potential. On the other hand, the least capable pupils inevitably show up poorly. Many of these less capable pupils, though they once have tried hard to learn, stop striving after they recognize that even with great effort they are still achieving only low marks.

In addition to these possible reactions toward being judged against the performance of classmates, there are numbers of others. For instance, some low-achieving pupils continually work hard in the face of very low-level success. Or there is the very bright, very diligent student who is not satisfied with being nearly the best or just barely the best in his class, so he is anxious about grades and works very hard. For the pupil perpetually at the top of his class, the result is undesirable if such competition could lead him to become insufferable in his supposed superiority.

The pupil's reaction to the system of comparing the pupil with a teacher-set standard depends upon many factors, including how high the teacher sets his standards. Student motivation may be affected differently in a class where it is known the teacher gives all high grades and in a class with a tradition of mostly low grades.

Comparing the pupil with his own apparent ability has become increasingly popular in the elementary schools to recognize the wide differences in ability of the children in a single classroom. Ideally, this system does not allow the bright student to become lazy. Instead, the bright pupil is held by the teacher to a high standard which is commensurate with his talent. The marks he achieves reflect how well he measures up to his potential. Likewise, the slow student's progress is measured against his own talent. If the slow learner is working well within the limitations of his abilities, he can achieve a satisfactory mark, although his performance is poor in comparison with his classmates'. This system, which is based upon a philosophy similar to that of the organizational practice of nongrading, is aimed at adjusting the mark to what realistically can be expected of the pupil.

In theory this method of grading is desirable, for it suits the mark to student ability. In practice, however, the potential advantages of the system of comparing the pupil with his own apparent ability are usually reduced by influences arising from tradition and human nature. For instance, school systems in the past have been geared to judging the pupil against his classmates or against the teacher's standard. With this tradition, it is often difficult for the bright pupil, who may not work to capacity, to accept a lower mark than that received by a slow, but diligent classmate. The slow student recognizes that he is not as capable as the bright pupil, and so the slow student may regard his own high mark with suspicion.

The teacher faces the problem of deciding the fair basis for judging each student. The teacher must usually make his estimate of the student's potential ability on the basis of either past performance or aptitude test scores. When past school performance is used as the base, the teacher may expect too little of a potentially bright pupil who has always worked below his capabilities. Likewise, the teacher

may expect too much of an intellectually limited but extremely hard-working pupil. When using aptitude test scores, the teacher is subject to the pitfalls of validity, reliability, and interpretation usually associated with diagnostic testing.

These, then, are some of the possible ways student motivation can be affected by the kinds of standards used in marking.

2. Parents' and friends' attitudes--Children usually try to do those things which will get them praise and approval from the people who are important to them, especially their parents and their friends. If parents deem marks to be important, they will usually encourage their children to work for high grades. Similarly, pressures may be exerted by peers if they are the kind who admire school success. Parental stress on grades, however, can reduce motivation if a child resents parental pressure. In this case the child, either purposely or subconsciously, does poorly in school so that the resulting poor marks will be a punishment. Excessive parental pressure may cause extreme frustration for the pupil who is already achieving to the limit of his capabilities.

In families which do not consider school success very important, the pupil feels little or no pressure from home or from peers to work for high grades. In fact, the student who applies himself and does well in school is often looked upon with suspicion by his peers.

Therefore, the extent to which grades motivate the pupil is governed partially by the attitudes toward grades exhibited by the people he considers important in his life.

3. The teacher's emphasis on marks--The teacher, by his daily actions, can focus more or less attention on marks. Some teachers, found less frequently in the elementary grades, use the promise or threat of a final mark as their prime motivating device.

The prospect of a final mark will stimulate some pupils to greater effort. This is probably more true in the secondary than in the elementary grades, because the older pupils have had several years of such conditioning and are adjusted to equating high marks with adult approval.

Placing great stress on the mark, however, is accompanied by some noteworthy dangers. Chief among these is that the pupil may focus on getting a grade rather than acquiring learning. The learning has thus become a concomitant of getting the desired mark. Such learning is often rote, not meaningful, and temporary; it was not sought as something worth while that could be used in the student's life. In addition, this striving only for a grade sometimes encourages the student to cheat or become unduly competitive.

The motivational functions of grades are very much subject to question, although, Wrinkle says, "...of all the functions marking and reporting practices are supposed to serve, they actually serve only one with any considerable degree of effectiveness--motivation" (121:33).

Psychologists and psychiatrists have noted the undesirable effect upon the mental health of many children caused by an overemphasis on marks. Educators are acquainted with the distress created during the grading process.

The question surrounding the use of marks for motivation lies in the way motivation is achieved. The process is often based on fear of failure, of humiliation, of loss of privilege--a negative form of motivation which research has demonstrated to be less helpful than the more positive means to the acquiring of learning quickly and permanently. Even when the search for a mark is based upon the positive desire for reward or accomplishment, care must be exercised so that the emphasis is cooperative rather than competitive.

If a student does not recognize value in what he is doing or is asked to do, the school may not be meeting the situation constructively by promoting learning activity through the motivational function of marks. A beneficial plan may be that of developing a curriculum which would involve real values, values which in turn would stimulate students to effective activity, and of improving instructional procedures which would make unnecessary the continued use of pressure devices.

Guidance

Few guidance workers are convinced that the grades which a student has obtained are in themselves sufficient evidence for making decisions about his abilities or about the courses he should take. In a school where the teachers are known to each other and the basis of their marks may be recognized by counselors, greater significance could be attached to the guidance function of marks, but even within a given school the basis of one teacher's marks may not be known by the others. Counselors, far from being impressed by the contributions that marks make to their work, are often more impressed by the problems they provide. Perhaps as many emotional and maladjustment difficulties of students are caused by grades as by any other factor.

The marks pupils receive at the end of a grading period are of some aid in planning current schoolwork. Day-by-day evaluations are

also important for helping the teacher design class work for the students' particular abilities. These daily appraisals analyze student skills into their specific components. School marks are quite useful to counselors at school levels where plans for differentiated education are being devised. Of course, numbers of other factors need to be considered in the educational counseling situation, such as test scores and student interests, but school marks should form part of the data used in guiding pupils' plans.

Administration

Students are often promoted, retained, or graduated on the basis of a grade-point average. American public education, however, is for all children, not for only a selected few. The school recognizes the wide range in abilities of children and tries to provide education appropriate for pupils of all levels. It is not the purpose of education to establish set standards of performance, and then to eliminate from school the children who have not met the standards as soon as some of their age-mates have. Failure for a child in his early school days may lead to frustration for his entire school career. The focus, then, is on each pupil's optimum growth, not on set standards.

Research studies provide guidelines for promotion policy at the elementary and junior high levels. Many of these studies compare the subsequent success of the slow student who has been retained with the success of the slow student who progressed with his classmates. Generally, it was found best to promote the slow student with his class, for he usually did better scholastically than if he were retained. But in some cases it is desirable for a student to be retained, or it is desirable for the gifted pupil to be accelerated. Many research studies do not recommend regular "automatic" promotion. Instead, the key question to ask at the elementary and junior high level is: "What will be best for the student?" A final mark which serves as a summary of test scores, ratings, and observations does not itself dictate promotion or retention. The decision about promotion is a careful weighing of all factors by the teacher in cooperation with parents, students, and school officials.

A composite evaluation of a pupil's achievement and growth should be available for his future teachers and for the administration. Whether or not the pupil remains in the same school system, information will be available about his progress in previous grades.

The school administrator or guidance officer is often asked to evaluate the competencies and potentialities of a school's graduates. Because more and more pupils are going to college, organized evaluations are needed for these pupils. Final marks are one of many aspects considered for college admission. Colleges are also interested in oral expression, laboratory skills, creativity, and other facets of a pupil's school career, such as the adequacy of his personal-social adjustment, his participation in extra-class activities, and the nature of his interests and goals.

Pupils who do not attend college are generally interested in employment. Because of the practice of social promotion, prospective employers often request specific information about a school's graduates. Competency in certain vocational skills is important for certain types of jobs. Employers also seek employees who can get along with people, who will assume responsibility, who possess leadership potential, and who have an attitude of cooperation.

The marks pupils receive at the end of a grading period are of some aid in planning current schoolwork, for these marks often reflect areas in which a pupil is strong or weak. Daily appraisals, however, are generally more useful guides than a mark, for the daily teacher judgments analyze student skills in specific components.

Marks provide records for the school. Such records help school officials in transferring pupils from one school to another, and they help teachers, administrators, and counselors plan future courses for the students.

Information

A pupil's knowledge of the results of his study is one of the conditions for effective learning. For instance, there appears to be value for the secondary-school pupil in having a comprehensive appraisal of his abilities in specific academic areas. The pupil needs information on his possible success in further study; this information will help him to decide on the course of studies he will pursue. Secondary-school marks are useful in predicting college marks; a number of studies show relationships represented by correlation coefficients well above 0.50 (101).

Interpretation of marks is important. Final marks should be interpreted in terms of the quality of the instruction and the general ability of the pupils in the class. In some schools where there may be a lack of student talent, high marks may give a distorted picture of pupils' achievement. The lower marks of some pupils may represent more achievement than A's in schools inadequately staffed and equipped. Likewise a final mark of C in an

honors course may represent greater achievement than an A in an average course. The value of the reported mark will be increased to the extent that the pupil is helped to understand its meaning. This implies that the teacher is prepared to defend its accuracy and to show its significance. Many teachers explain the meaning of their marks and other data before their pupils take the report home.

In order to promote cooperation between home and school, parents must understand the school's objectives. Some parents feel that academic achievement should be the only concern of the school. Many of these parents believe that certain kinds of academic achievement are more important than others. For example, the learning of dates, names, and places may seem more desirable than an understanding of the larger social issues involved. The mechanics of writing may overshadow the ideas expressed. A large vocabulary may seem more important than understanding what is read. A knowledge of facts is preferred to vocational training.

Parents are often perplexed when the school is concerned with group living, vocational training, and personal-social adjustment. When parents can understand and accept the school's objectives, they are in a position to support and supplement the educational program of their child. Reports formulated in terms of children's progress toward these objectives are informative and provide the basis for home-school cooperation.

Bases for Marking and Reporting

The bases for marking and reporting pupil progress are: objectives of the school, growth and achievement, and standards and ability to learn. Each of these bases has certain problems or limitations.

Objectives of the School

Effective instruction and subsequent evaluation are based on the educational objectives of the school. These objectives are generally stated in terms of desired pupil behavior. It is not possible to identify for all schools and for all times those educational objectives that are most important or to determine when they are to be emphasized. These objectives must be determined in terms of individual differences of pupils, teachers, and communities.

If evaluation is to be made in terms of all the general and specific educational objectives of the school, the marking and reporting become very complex. In the first place, the evaluation of some of the pupil's behavior will not be as adequate as one would like it to be. On the other hand, it may be more valuable than a paper-and-pencil test score that indicates only what the pupil says he will do. If one objective is to develop effective citizenship among pupils, statements such as, "He votes in all student elections for which he is eligible," "He assumes responsibility in group work," give some direction and provide a basis for evaluation. Secondly, it would be impractical to sample enough behavior of every pupil to obtain an adequate evaluation. This would be especially true for the secondary school, where a teacher meets many pupils every day. Thirdly, to report on each one of a great number of behavioral patterns would give the false impression that valid judgments can be made in every instance. A fourth problem arises in presenting such a report to parents. Its length and complexity mean that it cannot be presented easily and quickly.

The teacher, therefore, faces somewhat of a dilemma. On the one hand, he wants to base his marking and reporting on all pertinent educational objectives, whether general or specific. On the other hand, the list of these objectives is so long and, in some cases, the evidence concerning them so difficult to obtain, that he faces an almost impossible task. Several approaches make the reporting task more manageable. Rather than attempting to report pupil progress in terms of all educational objectives on each occasion that a report is given, the teacher may report only on a selected group; on a following occasion, he may report on a different group. Another approach is to report on each pupil on each occasion in terms of a group of basic objectives of special importance to the pupil and the school; each of these may be, in reality, a combination of related specific objectives.

Growth and Achievement

Reports may depict growth, achievement, or both. Growth and achievement are different concepts. Growth means change or gain. To interpret growth adequately, one must consider the individual, his ability, his background experiences, his present environmental stimulation, and so on. Achievement means the pupil's present status--what he knows or can do now. Achievement is generally compared to norms or to teacher's standards. Frequently, individual differences with respect to ability and prior experiences are not considered when evaluating present achievement.

Differences between the two concepts can be more clearly described through the use of an example. Let us assume that John and Bill have just entered the fifth grade. Each boy's achievement in reading comprehension at this particular time is represented by a point called prior achievement. Bill's prior achievement in reading comprehension exceeds that of John's by

considerable amount. If we examine the boys' achievement at the end of the academic year (final achievement), we find that Bill still surpasses John. Their growth in reading comprehension is represented by the difference between prior and final achievement. Both boys have grown with respect to this skill. But John has made a tremendous spurt. There are several interpretations that can be made, depending on other information that we have about these pupils. Suppose they have equal potential in reading. We might conclude that Bill is an over-achiever who could not be expected to make tremendous growth strides, whereas John may be achieving in terms of his ability. It is more likely that John was retarded in the earlier grades and through good teaching was able to make strides in the fifth grade. Bill, on the other hand, may or may not have made adequate growth during the year. Suppose, however, that Bill is a brighter boy than John; Bill's actual growth, then, is inadequate. Although Bill's final achievement may be the highest in the class, it is still unsatisfactory in terms of what he could do.

One can see from this illustration that both growth and achievement guide the teacher in reporting pupil progress. Should growth and final achievement be given equal or unequal emphasis in evaluation? What marks in reading comprehension should John and Bill receive? The answers to questions such as these depend in part on the grade level of the pupil. At no time is either growth or final achievement ignored by the teacher, yet the relative emphasis they receive shifts as the pupil climbs the educational ladder.

During the elementary-school years, the teacher is concerned with helping each individual achieve as much and as rapidly as possible. Studies in child development inform us that children mature at different rates; every child is different, and every child has environmental problems that affect learning and are peculiar to him. It is not always possible, therefore, to predict which pupils will eventually lead the group in specific areas of achievement. The elementary-school teacher is in a position to manipulate the school environment so that pupils are not thwarted and stunted in their growth before they reach a degree of maturity where they can realistically evaluate their own strengths and weaknesses. Very often, therefore, the elementary-school teacher emphasizes self-improvement--in other words, growth. Emphasis on growth stresses the child rather than subject matter; it can apply to all kinds of learning situations, academic and nonacademic. In the elementary school, growth is more important than final achievement.

Even in the secondary school, evaluation based on final achievement alone does not tell the whole story of the pupil's development. Achievement scores, for example, do not indicate whether the pupil is at a particular level because of lack of effort or lack of talent. A single achievement score gives no indication of whether or not there has been improvement. In the secondary school, when the pupil is looking forward to the roles he will eventually play in higher education and in professional life, the teacher must be concerned with standards of achievement as well as growth in evaluating his competency. The standards that students must meet in law schools, medical schools, and teachers colleges, for example, exist as much to protect society as they do to select qualified students. Competencies based on standards of achievement are therefore mandatory for effective teaching in the secondary school.

Standards and Ability To Learn

An issue that in some ways is similar to the growth versus achievement question is whether to evaluate pupils in terms of standards or in terms of ability to learn. Standards of achievement are generally thought of as the teacher's estimate of the level of achievement that a pupil must reach before he is judged to have done acceptable work. When a pupil is evaluated in terms of his ability to learn, his effort must be considered. All pupils, whether talented or not, will receive praiseworthy reports if they achieve as much as they possibly can under existing circumstances.

These two points of departure for evaluating pupils are not so different as one might surmise. Evaluating in terms of standards does not mean that individual differences among pupils are ignored, nor does evaluating in terms of ability to learn mean that standards have been abandoned. When standards are the criterion for evaluation, it does not follow that the talented pupils will be discouraged or even prevented from progressing above the passing level. Neither does it mean that the less able pupils will fail to experience success. When evaluation is based on a pupil's ability to learn, greater achievement is expected of the able pupil than is expected of his less able classmate.

In other words, whether instructing or evaluating, the teacher cannot disregard either standards or individual differences in ability to learn. Standards are multiple. Those for the superior pupil differ from those for the less able. Moreover, the level of the standards change from grade to grade, and to some degree, from curriculum area to curriculum area and from teacher to teacher. Nevertheless, there is a minimum level of achievement that a pupil must reach before he gains credit for a course or before he can be promoted to a higher grade level. If a pupil fails to

achieve at this level or a higher level because of inadequate ability, he should be moved to a classroom situation where educational experiences better suited to his ability are presented. Hence, the selection of appropriate classroom experiences is an important factor in providing for individual differences. Special provisions may need to be made for the elementary-school pupil. Guidance in the selection of appropriate courses is necessary for the secondary-school pupil.

Evidence of Pupil Progress

The specific evidence needed so that a pupil's progress can be accurately determined and reported must come from many evaluation instruments and techniques. Certainly data from paper-and-pencil tests, both informal and standardized, ranking and rating scales, and check lists should be used. In addition, the results of observation of class participation and pupil success in projects and reports are helpful.

Informal Tests

In common practice, the informal test is employed extensively as a partial basis for determining final marks and for making reports. Not only is it used for measuring the level of a pupil's knowledge and understanding, but it can also provide helpful diagnostic information. When well-constructed, the informal test is invaluable in directing the learning of pupils.

The informal test can be focused on the topics presently being discussed in the class. In reading, comprehension tests can give an indication of ability to select main ideas, to summarize, to determine meanings of unfamiliar words from context. Informal arithmetic tests can check skill in number combinations and processes, and accuracy and speed in working out these arithmetic examples. An analysis of the difficulties found by pupils at all levels in working word problems is possible by varied test items. Daily or weekly spelling tests reflect a pupil's progress in spelling and the legibility of his writing. In the content areas, knowledge of facts can be checked. Understanding of various skills and uses of equipment in science can be determined by informal testing. Such skills as map reading and knowledge and use of various kinds of reference materials can be evaluated in social studies classes.

Informal tests can also deal with learning difficulties that are of particular concern to the group. Teachers, of course, should exercise care that their tests do not place a premium upon verbalization and rote memorization of facts. If tests are not used correctly, they can set the learner for short-term retention. The pupil may be able to pass the quiz but fail to see the broader relationship of a unit of work.

Standardized Tests

The workmanship in the better standardized tests is generally superior to that found in the typical informal test. Great care and extensive trial administrations result in test items that are more likely to test the pupil in terms of the purposes for which the test is designed. In addition, there are norms by which the pupil's progress can be compared with national averages. In most instances, however, it would be unwise to base reports entirely on these test scores.

Standardized tests are based on objectives that educational authorities generally have agreed upon. There is no general consensus as to the emphasis these objectives should have or, in many instances, at what grade levels they form the base for instruction and evaluation. These are decisions that must be made by the local school and community. Therefore, to depend entirely upon the use of standardized tests in marking and reporting would presuppose that the teacher, the school, the pupils, and the community had established the same objectives and emphases as had the test makers. Only if this were true would the standardized test give a valid evaluation of the extent to which the class objectives had been reached.

Standardized tests can make a genuine contribution to the instructional program and to effective reporting. They provide a more accurate total picture of a pupil's strengths and weaknesses in subject-matter achievement than informal tests. Available norms also enable the pupil's achievement to be compared with that of pupils in schools other than his own. The diagnostic tests in particular can be helpful in determining the basis of a learning difficulty. Of course, the aptitude test scores indicate a pupil's ability to achieve. They provide criteria by which the adequacy of achievement can be judged. In addition to the achievement and aptitude tests, personal-social adjustment inventories can provide helpful data.

Class Participation

Class participation is the broad aspect of the pupil's participation in group activities rather than recitation per se. Evaluation should not be based on the number of times a pupil answers questions in class. There are some pupils who crave attention. Others may

be far better problem solvers and think more creatively than their more vocal peers. One can learn about pupils through the comments they make. Do they show that they have an understanding of the subject? What attitudes are reflected in their discussion? How does the class react to comments of individual pupils? Are ideas accepted or rejected because of the personality of the contributor rather than because of the essence of the remark?

Participation in class also involves getting along with others and playing one's role as a leader and a follower. Class observation is the only way in which certain personal-social aspects of the pupil's behavior can be interpreted. There is a problem when it comes to reporting this behavior objectively. Some of the subjectivity can be eliminated by evaluating in terms of educational objectives stated as observable behavior.

Procedure and Product Evaluation

Achievement of certain educational objectives cannot be evaluated effectively by paper-and-pencil instruments. Operating a microscope, writing a theme, delivering a speech, or preparing a food product are kinds of behavior related to some of these objectives. These performances are classified as procedures and products. Ranking and rating methods, check lists, and anecdotal records are generally used in evaluating procedures. Product scales as well as ranking and rating methods are employed in product evaluation. Sample products prepared by the pupil serve well as a means of explaining to him and his parents the degree of his achievement. Data yielded by procedure and product evaluation often require less explanation than do some test scores.

Projects and Reports

Some pupils are quite adept at learning facts and following prescribed directions and outlined procedures; but some important educational objectives represent higher intellectual abilities and are harder to evaluate with the conventional instruments. Such objectives include problem solving, creative and independent thinking, and the ability to locate and select pertinent information. Through the use of projects and reports, these objectives can be observed and evaluated. Projects and reports provide occasion for pupils to muster their inner resources apart from the artificial atmosphere of the paper-and-pencil test; they permit the pupil to attack a problem on his own. Furthermore, pupil interest can often be determined through the use of projects and reports; perhaps pupil effort, which has not been forthcoming in teacher-dominated activities, will be discernible in opportunities allowing greater freedom of action. Achievement in terms of these objectives is not easy to assess. Reliability may be low. If one feels that such activities should be included in the determination of final marks, then one will have to cope with the problems of inadequate degrees of validity and reliability.

Determining the Meaning of a Mark

As noted earlier, a variety of different symbols for marking are used by different school systems. Some schools use numbers, others letters. Some use percentage marks, others verbal descriptions. But it should be clear that, whatever the system, the symbols themselves have no inherent meaning. The meaning is assigned to a mark by the people who use it.

Within many school systems the staff has reached no really specific agreement about the meanings of the marks they use, so the mark given by one teacher does not mean the same at all as the identical mark given by another. It is most desirable within a school for the staff to establish as much agreement as possible concerning the meanings of the marks. If the mark is a letter or number grade intended to compare the pupil's progress with that of his classmates, the agreement can take the form of a description of the quality of work and the kind of pupil that is represented by each mark. For example, here is a description for the meaning of the mark of C in a junior high school:

> A pupil receives C when he:
>
> Is generally cooperative and reliable.
>
> Does quite acceptable work, but requires frequent guidance from the teacher, because he cannot work independently for any length of time.
>
> Gets along with classmates and teacher with little friction most of the time.
>
> Tries to do his assigned part in group work but does not take a leadership role or offer many fruitful ideas.
>
> Has only minimum interest in the subject, so does not pursue it beyond bare required work.
>
> Usually fulfills assignments.

This description is a general one, intended to be applied to a range of grades and a variety of kinds of classes. Such descriptions are even more useful if they are stated in a way that applies them more specifically to the objectives of a particular grade (such as seventh) and specific subject matter (such as social studies or mathematics).

In upper grades in which a student is compared with classmates, the school staff may not

create such descriptions as that above, but may define marks in terms of the quarter or half of the class the pupil falls into on the basis of the quality of his work. Here is one such description:

> The mark of 1 means: The student succeeds as well as the top 25 percent of his classmates.
>
> The mark of 2 means: The student succeeds as well as the middle 50 percent of his classmates. That is, his work is better than the lower quarter of the class, but not so effective as the top quarter of the class.
>
> The mark of 3 means: The student's work is of the same quality as the lowest 25 percent of his classmates.

Such descriptions in terms of quarters do not commit any particular percent of the class to fail. Whether any of the pupils in the bottom quarter of the class are retained in the grade depends on decisions concerning what will be best for each pupil in his individual case.

If, however, the mark is based on a comparison of the child with his own apparent abilities, descriptions of the meaning of marks will take a different form. For example, for intermediate-grade classes the marks might be defined in such terms as these:

> The H pupil: Always strives hard, always does his best at every task. One could not expect more progress for a person of his ability.
>
> The S pupil: Usually works up to his ability, but on some tasks does not do as well as he is capable of doing. Work is satisfactory, but might be improved.
>
> The L pupil: Usually seems content to perform at a level somewhat below his ability. Makes progress, but is likely to quit or reduce effort when he meets any difficulties.
>
> The U pupil: Makes little progress. Level of performance is far below capabilities. Needs much more effort or help in order to progress at a level equal to his potential.

These, then, are a few of the ways marks can be made more specific and understandable for the school staff.

Factors That Influence Marking

A mark represents the evaluation of the ability and accomplishments of one person by another. Marking is an extremely difficult process, and it is even further complicated by the pressures originating in the following sources: fellow teachers, pupils, parents, administrators, and future users of grades.

Fellow Teachers

School personnel are cognizant of the importance of favorable public and interpersonal relationships. Grading is a strong factor in determining these relationships. Teachers bring pressures to bear on other teachers because of their special interest in particular pupils, their interest in pupils participating in extracurricular activities under their direction, and their desire to create a favorable impression in and for the department. Teachers are often criticized for being either too strict or too lenient in their grading. These pressures may influence to a considerable extent the grades recorded by teachers, because they wish to have harmonious relationships with other faculty members. Indirectly, teachers are influenced by other teachers in that examination of pupil cumulative records (halo effect) indicates wide divergence between their opinions of students and the opinions held by other teachers.

Parental Pressures

The pressures exerted by parents, either directly or indirectly, are many and varied; i.e., anger, loss of face and social prestige, "apple polishing," comparison of siblings, apathy, and parents' holding to unrealistic goals for their children. While many other examples could be mentioned, these illustrate the influences that affect the teacher's grading.

School Administrator Pressures

School administrators can influence the grading of students by teachers. Some administrators desire to maintain a high ratio of pupils going on to colleges and universities. They are also interested in a good public relations program and in presenting to the patrons an excellent image of the school.

Another consideration is that to alleviate the crowding of schools in times of classroom shortage, some administrators have encouraged teachers to pass all possible students and keep them moving through school.

Teacher promotions are sometimes based on the success of the teacher in getting pupils to rate high in state-wide testing programs, the types of grades they give to the pupils, and the amount of trouble they get into with the parents.

To a limited extent, grades are also used for planning the curriculum. If it is found that pupils do not understand certain subjects too well and receive low marks, administrators sometimes drop these courses from the curriculum in order to maintain a higher level for all pupils. Sometimes, courses are changed or modified on the basis of the understanding demonstrated by pupils.

Pressures from Future Users

Users of grades are aware that good school marks are one of the better predictors of a capable employee in the world of work, or of success in higher education. Teachers know that much significance is attached to good grades, and they do not want to hinder the achievement of pupils' vocational and educational goals. Although these pressures are indirect, teachers are cognizant of their influence in considering the pupils' marks.

Pupil Pressure

Students definitely wield an influence with the classroom teacher. For instance, there are studies which show that girls get school marks that are consistently higher than grades received by boys. Another personal factor that influences the teacher's grading is that of the pupil's I.Q.

By and large, teachers would like to be friendly with their students, because it is axiomatic that it is more pleasant to teach those who like you and to like those you teach. Pupils, however, realize this and make it a goal to get on the proper side of the teacher. Even though "apple polishing" is prevalent, most pupils will violently reject anyone who consistently behaves in this fashion. It is probably true that the grading process has cast suspicion on the student who is genuinely interested in the importance of his ideas and is actually studying because he wants to learn.

Students also attempt to arouse the sympathy of teachers toward their need for good grades because of the students' desire to participate in various types of school activities. Too, there is a need for a respectable record which influences future education and vocational planning and goals.

The fact that grades have been good predictors of success indicates that teachers have been doing a competent job of grading. An even higher degree of competency could be obtained if all teachers were truly aware of the multiplicity of factors that influence grading.

Summary

An attempt has been made in this section to relate some of the multiplicity of factors which prevail in marking and reporting pupil progress within school systems. Specifically, some of the major criticisms of marking and reporting as they affect the learning process have been highlighted. A list of the arguments condemning school marks include these ideas:

1. Marks are variable, subjective, contaminated, even capricious.
2. Marks create a condition of unfair competition.
3. Marks reflect an aristocratic rather than democratic attitude.
4. Marks preoccupy students and their parents.
5. Marking practices deny the psychological principle of individual differences.
6. Marking practices tend to influence teaching in the direction of memorization and regurgitation at the expense of concept formation and creativity.
7. Marking practices encourage student dependence.
8. Marks frequently have an emotional impact that is at variance with good mental health practices.

In reporting current progress, should the teacher compare the pupil with his own past achievement and ability or with his classmates? Or should he be compared both with himself and his classmates? These questions pose a basic problem in educational philosophy that plagues the teacher each time a pupil is tested, rated, judged, or marked.

Defenders of the viewpoint that children's marks should reflect their relative standing in a class say that:

1. Children must learn realistically what their abilities are.
2. Children must learn to recognize their areas of low ability, where they receive low or failing marks, and where they need additional work.
3. If a student's records are inspected by a college or an employer following his school career, marks comparing him with his classmates will give the college or employer a better estimate of his abilities.

4. If a student realizes in school what his abilities are, he will be realistic and not expect to be successful in areas in which he has little or no ability.

Defenders of the viewpoint that children's marks should reflect the progress they have made in relation to their individual abilities say that:

1. The elementary school does not have the function of comparing children with each other or eliminating the less apt, but it operates to provide opportunities for each to learn to the best of his ability, despite what that ability is.

2. Constant defeats for a child, as shown by consistently low marks when he is compared with others, are damaging to his personality and do not give him the supposed realistic and healthy attitude toward his abilities that proponents of competitive grading claim.

3. When individuals, as older youths or adults, are being trained for particular vocations, comparative marking may be appropriate to distinguish the more able from the less able. But vocational selection is not the function of an elementary school in a democracy. Instead, the elementary school should provide for all ability levels.

Other educators take a compromise view. They desire that children be compared both with their own apparent abilities and with the achievement of their classmates.

Which of these points of view is the correct one or the best one is a philosophical matter that each school or school system must decide for itself. All three philosophies are reflected in report forms from various parts of the country. In recent years, as more information about individual differences among children and mental hygiene principles has been developed, there has been a noticeable trend toward comparing the child with himself, at least in the lower grades.

METHODS OF REPORTING

IN ATTEMPTING TO IMPROVE marking and reporting, schools throughout the United States have developed a variety of methods and forms. It would not be possible to decide which of these methods and forms is the best. Each has advantages and disadvantages. Some are better suited to a particular kind of school or community than others. A description of some of these methods and forms will demonstrate their values and limitations. Such descriptions may provide suggestions for teachers and school systems which are developing reporting methods to accurately tell pupils, parents, and the school administration of the pupils' progress.

Typical procedures for marking and reporting are letter-number systems, check lists, anecdotal records, behavior descriptions, correspondence with parents, teacher-parent conferences, and self-evaluation. It is important that these methods be critically evaluated in terms of purposes, bases, and available sources of information for reporting. Whichever is used must be established as the best possible system under existing circumstances.

Letter-Number Systems

The letter-number system has been and still is a popular method of reporting pupil progress. In essence it is an effort to summarize a great variety of information about such pupil characteristics as growth, final achievement, effort, ability, and general deportment. The summary is given the form of a letter, usually A through E or F, or a number, usually 100 through 0. Such a letter or number is usually reported for each subject area each report period. These marks are most difficult to assign, for two important reasons: (a) the teacher's uncertainty as to whether an appropriate amount of credit is being given to effort, growth, and final achievement for each educational objective considered; and (b) since the mark represents pupil progress in terms of a combination of objectives, the teacher's uncertainty as to whether each objective within the combination is being appropriately weighted.

The report card used to record the letters or numbers assigned can take many different forms. For example, rather than identifying the five points in marking system by the letters A through F, the letters H (with honor), M (with merit), S (satisfactory), U (unsatisfactory), and F (failing) may be used.

Dual Marking Systems

An interesting variation of the traditional letter-number system is the use of not one but two marks for each subject. The first mark represents the pupil's final achievement in relation to the teacher's standards. The second often represents the amount of effort put forth by the pupil; in some school systems it represents pupil growth in the area of achievement listed.

An illustration of the idea of dual marking follows. One report card uses a five-point letter system for recording the pupil's final achievement. A three-point letter system is used for recording the pupil's effort: H if the pupil exceeds what is expected of him, S if his level of work corresponds approximately to his level of ability, and U if the pupil is capable of better work. These three letters are called "ratings of achievement based on ability." In another report card, intended for use in the seventh and eight grades, the pupil's final achievement is shown in terms of percentages (75 percent is the passing mark) and again a three-point letter system is used to evaluate effort: E for excellent, S for satisfactory, and U for unsatisfactory. Effort is described as suitable daily preparations, attention in class, and participation in class.

Pupils evaluated according to these systems might receive marks such as A/S in the first case, and 91/S in the second. In both instances the final achievement is considered to be high and the effort normal. Marks such as C/H or 83/E mean that the pupil is average in terms of final achievement, but he is working well up to his capacity. His effort is superior.

Clearly, the success of these dual marking systems rests upon the teacher's ability to evaluate final achievement and effort by means of a letter or number. Teachers sometimes find the evaluation of effort particularly troublesome. To judge a pupil's effort in a satisfactory manner means that more accurate evaluation techniques must be available. On the other hand, if the second mark represents

growth, an extensive pre-testing program must be established so that the amount of prior achievement is known before instruction begins.

The dual marking system can be expected to provide a clearer picture of the pupil's progress than does the single mark so often used, but they are still an oversimplified picture of the teacher's evaluation. Each subject-matter area listed on the report card encompasses not one but many educational objectives. With suitable instruments and sufficient time the pupil's final achievement and effort or growth could be determined and reported in terms of each of these educational objectives.

Relative Versus Absolute Standards for Marking

The lack of an absolute zero for achievement test scores makes it impossible to assign marks on an absolute scale or standard. If a pupil correctly responds to all test items on a test, it does not necessarily mean that he knows 100 percent of what is to be learned in the subject; a mark of 50 percent does not mean that he has learned 50 percent of the subject matter. Test items are but a sampling of what is learned. Also, whether a pupil receives credit for knowing a particular concept depends upon the manner in which his understanding of that concept is measured with a particular item. Even a question about a difficult concept may be answered correctly by almost all pupils if only a superficial understanding of the concept is required. In other words, item difficulty often depends upon how a concept is measured as much as on what is measured.

Actually teachers who convert raw scores or percentages directly into final marks usually adjust these raw scores or percentages near the end of the term by giving bonus points, by constructing an easy or difficult achievement test as the situation may warrant, and so forth, until the raw scores or percentages would convert to a distribution of final marks which was largely predetermined. This distribution of marks involves an arbitrary decision. It may be determined by school policy, the ability and vocational plans of the pupils, the presence of gaps in the distribution of composite scores, etc. For example, in a high ability class a teacher might give 30 percent A's, 40 percent B's, 20 percent C's, and 10 percent D's. Rarely are there as many F's as there are A's; the median final mark in many classes is often B rather than C.

Regardless of the marking system adopted, it is important that the marks present a clearly defined message to pupils, parents, employers, college admission officers, and others. Clarity of meaning is enhanced when there is agreement among school personnel on the criteria by which the percentage of pupils in a given class will be assigned a particular mark.

To illustrate how final marks may be assigned unfairly, consider three classes in English taught by different teachers. Classes 1 and 2 are populated with college-bound pupils of comparable ability; Class 3 has pupils of lesser ability. The teacher of Class 1 assigns 60 percent A's and B's; the teacher of Class 2, who is less liberal, gives only 40 percent A's and B's. Consequently, a pupil whose English achievement is typical of college-bound pupils would probably receive a B if he were in Class 1, a C if he were in Class 2, and an A if he were a member of Class 3. In the latter class it is likely that he would be one of the better students.

One way of rectifying this situation is to obtain some common and relevant measure by which the three English classes could be compared. Scores from a scholastic aptitude test or from standardized or informal English achievement tests with acceptable degrees of content validity are measures that might be used. Suppose it is decided that 15 percent of the pupils will receive A's. Then the cut-off score on the common measure would be determined, above which only 15 percent of the pupils score. The number of pupils in a particular class who surpass this cut-off score provides the approximate number of A's to be awarded to members of that class. This is likewise true of the percentages of B's, C's, and so forth. Note that the final mark a particular pupil receives will depend on his performance. It does not necessarily follow that a pupil who made a scholastic aptitude test score in the B-range will receive a final mark of B. Some pupils will receive higher final marks than those with higher scholastic aptitude scores because of class work and performance.

Weighting Data

One of the problems the teacher encounters in giving a final mark is weighting the data from various sources, such as informal quizzes, final examinations, reports, and so forth, in order to obtain a valid composite score for ranking pupils. There is no consensus as to the emphasis each of these types of data should receive in the total evaluation. Some teachers maintain that the final examination should count far more than the quizzes and other tests given during the course. They point out that the final examination measures long-term retention and ability to organize and deal with large units of subject matter. They conclude, therefore, that it is logical for this test score to be given more weight. Others are quick to retort that it is quite unfair to place a pupil in a position where such a large

proportion of his final evaluation should be determined on a given day at a specified time and within an interval of at most several hours. They also object to the limiting features of many final examinations that preclude measurement in terms of many important educational objectives. This is a very important consideration. Weighting various data must be determined in terms of the educational objectives of the specific grade level or class. Those data that reflect pupil progress in terms of the most important objectives must be given greater weight when computing a composite mark than those reflecting pupil behavior in terms of less important objectives. This is true irrespective of whether the data come from a final examination or from any other evaluative instrument.

There is a precaution that must be taken in determining a composite mark no matter what weightings are chosen. Suppose we wish to base one-fifth of the final mark for a course on class reports, one-fifth on daily assignments, one-fifth on quizzes and unit tests, and two-fifths on the final examination. Inspection of the ranges and standard deviations of these measures in the following table reveals a definite lack of uniformity.

Source of points	Nominal weight	Range	Standard deviation
Class reports	1	26	5.1
Daily assignments .	1	46	8.6
Quizzes and unit tests	1	110	20.8
Final examination .	2	90	16.4

If one hopes to maintain the weighting scheme originally chosen, one must take into consideration these differences in variability. A failure to do this will result in inequities. To illustrate this point, let us suppose that Mary made a total of 55 points for class reports, the highest number of points for this category. However, on quizzes and unit tests her score of points was the lowest, 30 points. For daily assignments and the final examination she earned 61 and 91, respectively. Frank, on the other hand, did the poorest of anyone in the group on class reports; his score was 29. On the quizzes and unit tests he had a high score of 140. It happens that he also made the same scores as Mary on daily assignments and the final examination, namely 61 and 91.

If one weights Mary's and Frank's scores for each source and adds them to obtain a composite, we have the following:

Mary	Frank
55x1 = 55	29x1 = 29
61x1 = 61	61x1 = 61
30x1 = 30	140x1 = 140
91x2 = 182	91x2 = 182
Total 328	Total 412

Note that although class reports and quizzes and unit tests are to have the same weight, Mary is penalized because of the lesser variability of scores of class reports when compared with that of quizzes and unit tests. In other words, class-report scores with a relatively low standard deviation (σ=5.1) have less influence than scores of quizzes and unit tests (σ=20.8) in determining the class rankings on the composite score.

A procedure for avoiding errors like the one above would be to convert the raw scores into standard scores so that the variability of the scores for each category would be the same. On the basis of this distribution the teacher may then assign final marks.

Reliability and Validity

The reliability of scores from paper-and-pencil tests is sometimes not as great as one wishes it to be. Even lower, customarily, is the reliability of data from a teacher's observations of a pupil's procedures and products and his personal-social adjustment. Yet these are the pieces of information used to determine final marks. Consequently, the reliability of final marks is oftentimes less than desired.

Such evidence as is available suggests that the usual reliability coefficients of semester marks may be as high as 0.70 to 0.90. This means that in many if not all cases it is difficult to defend the practice of interpreting such differences as those between an 83 and an 84 or a 91 and a 92. For that matter, the difference between C+ and a B- may be due to chance only.

Evidence concerning the validity of final marks is also limited. They seem to be quite valid as a measure of mastery of subject matter by the pupil, the correlation coefficient estimated to be 0.70 or possibly higher. This estimate was arrived at by summarizing the results of a number of studies concerning the correlation between the final marks in question and (a) other marks in the same subject matter area, (b) test scores from appropriate standardized tests, and (c) the pupils' estimates of the marks they deserved.

The predictive validity of final marks has been investigated with respect to a number of different criteria, such as college entrance examinations, college marks, economic success, and success on the job. The correlation coefficients vary a great deal. The most widely investigated use of secondary-school final marks is that of predicting academic success in college. The correlation coefficients are often as high as 0.50 and seldom higher than 0.70. It has been repeatedly shown that final marks in secondary school are one of the best means

of predicting college success, whether used alone or as part of a prediction battery.

Interpretation

The typical letter or number system of marking is based primarily on final achievement. When an attempt has also been made to include an evaluation of effort and other personality traits in a single mark, confusion results. Interpretation becomes almost impossible. There is no way to determine whether a low mark results from lack of achievement or lack of effort. Even when used to report final achievement alone, it is very difficult to make a valid interpretation. What does a B in general science mean? Does it show that the pupil did A work in quizzes, C work in laboratory, B work in class participation? Or, more important still, does it mean that the pupil achieved at an A level with respect to one educational objective, a B level with respect to another, and a C level with respect to a third? One has no way of knowing. The marks are ambiguous. Therefore, when letter or numerical marks are used, there must be additional data in the report that provide diagnostic information.

The mark is also influenced by other factors, one of which is the nature of the pupil population. Too often pupils who have been graduated magna cum laude from their high schools find to their dismay that they do not have exceptional ability when they face competition in the university. Another factor is the philosophy of the individual teacher who does the marking. Some teachers give consistently high marks while others pride themselves in never being so easy as to award an A or a 95. A pupil may therefore find it difficult to identify a criterion by which he can judge the value of his marks.

Marks, as they are now generally assigned, are far from being as meaningful as many people think that they are. They are being overinterpreted; they reveal far less about the pupil than commonly supposed, and their meaning is often ambiguous. Hence, it is often recommended that standardized test scores be employed whenever possible to supplement final marks. Standardized test scores may provide a better opportunity to interpret pupils' potentialities for future academic work.

Pass-Fail Systems

The pass-fail system is considered by some to have advantages over the traditional A-F system. The first and most widely stated of these advantages is that it promotes healthy exploration in courses which might otherwise be avoided for fear of endangering the grade-point average. A second is that the system allows students to apportion their study time as they wish, rather than expend most of their efforts on courses which prove the most difficult and hence, pose the most serious threat of an undesirable grade. A third advantage of the system is that it removes the burden of evaluation from creative students, who may be penalized by the traditional A-F system.

One possible consequence of the pass-fail system is not considered desirable. This attribute of the pass-fail system is that it permits a pupil to earn course credit without being meaningfully evaluated. Students high in what Atkinson (7) and Atkinson and Feather (8) called "fear of failure," therefore, might elect the pass-fail option, regardless of their interest in the actual subject matter, simply to avoid an undesirable test experience and the concomitant threat of possible failure. If this were true, one of the chief purposes of the pass-fail system, to encourage intellectual curiosity, would be negated.

Check Lists and Rating Scales

Reports of academic achievement are often accompanied by a check list of personality traits and attitudes. The teacher may generally place a check in any one of several categories, such as "unsatisfactory," "satisfactory," or "improving." If the phrasing in these check lists is carefully worded, the instrument can be economical in terms of the time and effort required by the teacher.

Care should be exercised so that reporting does not become a stereotyped procedure. In the past many pupils have been checked "satisfactory" on every characteristic from kindergarten through the sixth grade. The "halo effect" seems to prevent some teachers from giving a realistic evaluation of the pupil. This can be eliminated if the teacher bases his report on adequate data, carefully compiled and critically interpreted. To make the check list more individualized, it may be constructed so that space is available for comments. This provides an opportunity to clarify the pupil's problems.

At the secondary level, check lists are generally used to supplement numerical or letter marks. Not only can they then provide information not discernible from the mark alone, but they often help to clarify the mark itself. On the other hand, check lists are sometimes the only report given in the elementary school.

Check lists and rating scales may be harmful since research has shown that among their limitations they (a) suggest that certain characteristics are equally desirable in equal amounts for all individuals at all times;

(b) encourage generalization about a pupil's characteristics beyond what was actually observed by the rater; (c) encourage the making of comparisons of pupils who are quite different and who have had unequal environmental opportunities; (d) assume that teachers can observe behavior, such as cooperation, sort it into units on a scale, and allot values to it; (e) suffer from "halo effect," that is, the teacher who rates a pupil high in one characteristic tends to rate him high in others or vice versa; and (f) usually suffer from inadequate definition of the terms to be rated so that what is satisfactory to one person may be unsatisfactory to another.

The basic difficulty with these methods lies in the judgment aspects of rating. In accepting a rating scale, one also accepts the philosophy of the person who constructed it. Research on individual differences suggests that it is unrealistic and unwise to try to place pupils on the same scale without considering their unique circumstances and situations.

Anecdotal Records

The anecdotal record of observing, interpreting, and reporting pupil progress has received a great deal of attention during the past two decades. This method of reporting can be of value when used properly and with caution. The anecdotes in a record are descriptive accounts of episodes or events in the daily life of the pupil, with some interpretation of their significance in his development. All classroom teachers observe these events and, unless an attempt is made to record them, they must depend upon memory to evaluate a pupil's growth. It is true, also, that if incidents are recorded and interpreted, the one making the record is more likely to be diligent in observation and more serious in his efforts to understand a pupil's behavior.

In practice, anecdotal records have been concerned more with social relationships than with subject-matter accomplishments; they can be of value in both areas. Thus, observation of a pupil in the classroom may reveal vigor or lassitude of response, variation from usual behavior under specific stimulation, a tendency to go beyond minimum requirements, attempts to improvise, reactions to authority, and relative degrees of zeal or apathy in response to various activities.

People who have appraised the written anecdotal record method are usually in favor of it, but the claims to objectivity in the method are open to question. The choice of a word involves a good deal of subjectivity. For instance, if a classroom teacher reports only that Jimmy "walked down the hall," it is said that the statement is objective. Another observer of Jimmy under the same circumstances might well have said he dawdled down the hall, still another that he strode and still another that he hurried. Research into the validity of the testimony of observers and the data obtained by specialists in word meaning raises considerable doubt as to the objectivity of observational techniques, even when they are used by trained specialists.

Advocates of the anecdotal record recommend that many records be made so that patterns of characteristic behavior may evolve. Caution should be exercised, however, so that school personnel are not alienated from a good evaluating and reporting program by insistence upon a large number of anecdotes on a specific number of pupils in prescribed form at a given time. Rather, the teacher should be asked to report behavior which is consistent with (or significantly different from) the activities of the pupil as he has observed and understood him. A workable plan includes also personnel to do the clerical work and the summarizing of the anecdotes.

Behavior Description Method

The method of behavior description can provide valuable evidence of pupils' development in the areas usually covered by check lists and anecdotal records. The entries in this behavior description method resemble those in check lists and rating scales, but there is a basic difference--the emphasis on the importance of individuality and the importance of individual differences. First, those who plan to use the method must prepare carefully defined descriptions of the pupil's behavior (sometimes similar to those used in check lists and rating scales). Teachers, and others with sufficient opportunity to observe the pupil, place symbols indicating their relationship to him beside the description which best fits the pupil.

The difference between this descriptive procedure and rating is just that the describers try to summarize what has been observed while raters attempt to judge the quality of the observed behavior. There is no implication in the behavior description method that any particular kind of behavior is best for any one pupil at a particular time. The technique admits the fact that a child's behavior may vary in different situations and under changing influences. Thus, though each reporter makes a correct description of what he observes, the reports about an individual may differ greatly at any given time. The plan allows for the possibility that differences in the descriptions of various observers may be as significant as the differences they report. It must be emphasized that there is no implication of goodness or badness in the use of the term behavior.

Instead of requiring a perfunctory rating of personality twice a year, a practice that teachers dislike and if possible avoid, the behavior description method purposes that teachers be encouraged to make continuous observations of their pupils with respect to the defined characteristics and to record their descriptions at such times as are decided upon. Duplicated sheets of the definitions of characteristics are furnished to the teachers so that they can make their descriptions of the pupils with the definitions before them, and without being influenced by each other's observations. The descriptions can be entered on sheets of class lists with the characteristics used as headings across the top of the page. Abbreviations and numbers for types make such a form simple to prepare. The descriptions are transferred from the class list to the central record card, thus making a picture of the pupil as seen by all his teachers. When significant notes accompany a description, they can be entered on the record card beside the definitions. If teachers study the form and the definitions of behavior at the beginning of the school year and agree to make the descriptions upon the basis of carefully considered evidence, the descriptions are likely to be valid.

Correspondence with Parents

Communications sent to parents may range from a little notation on a report card to a long letter discussing many aspects of their child's school experiences and growth. Both pupil development and subject-matter achievement can be emphasized. With an adequate cumulative record, the teacher has evidence to interpret patterns of development. There is an opportunity to discuss problems and emphasize factors pertinent to the individual pupil. If there are phases of his work that would profit from help at home, this fact can be explained to parents. There is also provision for two-way communication in that parental comments are requested. It is possible that the interchange of information may also lead to some fruitful conferences between teachers and parents.

One of the serious weaknesses of letters to parents can be their generalized and stereotyped nature. Too often they give the appearance of having been run off an assembly line, with little variation in wording to relate to a particular pupil.

Communicating to parents through written letters is not easy. For example, among other comments to a parent is the statement, "Your child is nervous." There may be an immediate reaction by the parent; he becomes concerned about the mental health of his child. The teacher had not meant to convey this idea at all. In other instances letters may irritate the parent or fail to express what the teacher wishes.

A capable teacher who writes lucidly can create an interesting and very useful letter for parents. However, some teachers either do not express themselves well in writing or do not keep adequate evaluation data to form a specific report of the pupil's progress. Because serious trouble may result from poorly written letters, many principals supervise this type of reporting quite closely. Some schools have provided outlines or suggestions for improving this correspondence.

It is obvious that to compose a thoughtful and helpful letter, a great deal of time is required. The teacher is faced with an enormous task if he has to report on 30 or 40 pupils. Some elementary schools have abandoned this means of reporting for this reason. It is impractical in most secondary school situations.

The problem can be somewhat alleviated if reports of all pupils are not sent out at the same time. Reporting can be staggered, thus spreading the load. Also the number of letters for each pupil can be limited; it is better to have one good letter than three poor ones. Other types of reports can be utilized for the other reporting periods. It is also possible that all pupils will not be reported by letter the same number of times. Some can profit from more numerous reports, whereas others will not necessarily benefit from as many.

As mentioned earlier, many schools do not rely solely upon letters to parents for the complete report. Instead, they use a printed report-card form which includes, in addition to marks on various skills and traits, space for teacher comments. In some cases this space is of considerable size, and a short letter home is generally expected. When writing such letters or comments, it is especially helpful for the teacher to have sufficient data about each pupil's specific behavior to create an individualized report.

Teacher-Parent Conferences

The conference method of reporting is most generally used in the elementary school. It has great potential for providing more information and better understanding between home and school, for misunderstanding can be more readily eliminated than in a written communication, and either conferee can raise questions for information or clarification. Of course, there are optimum conditions for conferences. Certainly, they should not always be called when a special problem arises; rapport cannot be at its best if this is the only basis for meeting. Rather, they should be planned

periodically to serve as a regular report on the pupil's progress.

Sufficient time must be available for an effective conference. To attempt to rush through an interview in order to be prompt for the next is unfair to both parent and teacher. Certainly, conferences are just as much a part of a good instructional program as is teaching the multiplication tables. Therefore, the teacher should be provided with the time to prepare and hold conferences during his working schedule. Some schools dismiss class for half-day periods; others provide substitutes to relieve the teacher of classroom responsibilities. Because a great deal of time is involved in this kind of reporting, some schools arrange only one conference per pupil during the school year, the reporting at other times being done by other methods.

Conference Preparation

Basic to an effective conference is preparation. To assist in this respect some school systems orient their teachers in the conference technique by distributing to them a bulletin describing the conference purpose, policy, records to be kept, time, and preparation. These bulletins typically contain numerous illustrations that suggest ideas for conference preparation.

Some schools have organized workshops and various kinds of training sessions for teachers. Role playing, in which teachers assume the role of both parent and teacher, is helpful. Writing a script of a hypothetical conference and inviting criticism from colleagues develops an insight into conference skill. Bulletins outlining suggestions including "do's" and "don'ts" are often issued.

It is helpful to anticipate some of the questions parents may ask in a conference. A list of such inquiries and suggested answers should be included in conference bulletins for teachers.

Whenever possible, the parents should be involved so that they feel they are helping their child to succeed. Too often parents have been told in answer to queries about their children's difficulties that everything is going to be all right. They are advised not to concern themselves with instructional problems. Nothing is more frustrating to the interested and intelligent parent than this approach; also, an excellent source of help for the child is blocked.

In addition to skill, the teacher's attitude about the conference to a large degree determines its success. He has to accept a parent and be able to look objectively at the parent's problems and their relationship to the child's welfare. Resenting the parent for his treatment of the child is more likely to reinforce the parental attitude than to change it.

Specific Information

Always have specific information available for each pupil. Dated and illustrative examples of daily work, quizzes, and anecdotes can mean the difference between a generally unproductive, even though friendly, visit and a purposeful, helpful conference.

Rather than comparing elementary-school pupils with each other, discuss the individual child's progress in terms of his past achievement, and his own strengths and weaknesses. Try to be constructive and positive in discussing pupils. There is a great deal a teacher can report on a child's progress if he has the data on which to draw. Along with information about a pupil's academic work, data on his health, tardiness, attendance, and personal-social adjustment should be reported. It is important to relate these data to his scholastic achievement, for they are not divorced from the total picture.

Limitations

Conferences have several important limitations, one of which is the time factor. Yet, as in the case of letters to parents, this time can be very important in developing learning readiness. Education of parents may be as important as working with his child in the classroom. Teachers also find that the information gathered and the thinking done in preparing for the conference may be very useful in class instruction. Often teachers are forced to collect information for a conference that they should already know.

Another limitation sometimes mentioned is that parents will not come to conferences. A great deal of the blame for this situation can be attributed to the school administration and to those teachers who are ill trained in the conference technique and who enter the conference with inadequate data. The effective administrator will involve the community leaders as well as his staff in affairs of the school. He is concerned about the community's understanding of the purpose of the conferences. Of course, he and his staff need to exercise the greatest care so that the conferences prove worthwhile and do not degenerate into sheer generalizing and passing of pleasantries. An inadequately trained or unprepared teacher can do more harm than good, for the parent can be antagonized and refuse to cooperate readily in the future. In such cases, the usefulness of the conference-type reporting is greatly reduced.

A third limitation is that the conference technique is difficult to use in the secondary school. It is true that most of the experience with conferences has been in the elementary school. Moreover, the consensus of many writers is that secondary-school pupils do not want their parents to participate in formal teacher-parent conferences. One study of 248 junior high-school pupils does not support this conclusion (90). Forty-four percent of the students favored teacher-parent conferences, 30 percent favored teacher-parent-pupil conferences, and only 18 percent wanted to exclude the parent.

No one would maintain that the child ceases to grow after he leaves the sixth grade or that parents have no place in his educational program. As the educational level of parents increases over the years, there is likely to be as great concern about pupil growth in the secondary school as there is now in the elementary school. In fact, there is some indication that this trend is developing. Nonetheless, there is a very practical problem that must be solved before there can be extensive conferences at the secondary-school level. The large teacher-pupil ratio will prevent conferences with parents of all pupils during the school year. There must be a basis for selection in terms of the teacher's need for better home understanding and parental cooperation in order to help some pupils. The guidance staff could alleviate the teacher's load by conferring with some parents.

Self-Evaluation

One aspect of the total evaluation program that is being emphasized more and more is the need for pupils to evaluate themselves. It is pointed out that self-evaluation might be used as a means of reporting pupil progress. After reviewing the literature, Russell (96) concluded that self-evaluation in school bears a positive but slight relationship to two other criteria of adjustment, teacher and peer ratings. Relative to self-ratings in academic achievement he further reported that elementary-school children who have had no special instruction do not appraise their academic progress as their teachers or peers do, although they are fairly consistent in overestimating or underestimating their achievements.

Russell also finds that socioeconomic class and a pupil's level of aspiration affect achievement as well as the opinions of his teachers and peers concerning this achievement. A few studies claim that teachers tend to approve middle-class standards of behavior because of their own similar background.

There is certainly insufficient evidence to support self-evaluation as the sole device for marking and reporting system. Rather, teachers should guide pupils in learning how to evaluate themselves. This is an important educational objective, and one helpful way of achieving it is to discuss a pupil's progress with him. The pupil should be able to interpret his achievement as it relates to his strengths and weaknesses. The teacher should help him to understand the various reports sent to his parents.

Summary

Common methods of marking and reporting are the letter and number system, rating scales and check lists, correspondence to parents, and the teacher-parent conferences. The letter and number system has a very definite weakness; it can tell only part of the story. On the other hand, correspondence with the parents and teacher-parent conferences are as broad in their evaluation of all aspects of the pupil's growth as the teacher cares to make them. Their limitations lie in the difficulty of effective communication and in the time required.

Increasingly, schools throughout the country are revising their report cards so that they are more detailed and are stated in terms of the skills children show. These newer cards or progress reports are becoming the chief means of informing parents of their children's growth toward the school's goals.

In the past there has been a tendency for the same report-card form to be used throughout the entire school, or at least from the kindergarten through the eighth grade. It contained a list of general subject-matter areas common to all levels. However, when the move was made to reporting more specific behaviors, it was realized that the goals in the lower grades were not the same as those in the upper grades. This has resulted in the development of separate report cards to fit the particular programs at the different grade levels. As a result, parents are receiving more information that makes considerably more sense. They know precisely the skills and activities developed in the school and their children's progress toward each of these.

Since the report card is often a main link between school and home, it is important for the card to be an effective instrument of public relations. The idea that school is interesting and that reports to the home are nothing to be feared appears to be conveyed by many cards being developed in all parts of the country. It is common for reports to be in multi-colored and, especially in the elementary grades, to have a cartoon or sketch on the front. Numerous report forms show the use of modern typographical design, and attractive colors make the progress report a pleasant appearing folder. Increasingly the

message to parents is friendly and mature. All this appears to bring the school and home closer together for the pupil's benefit.

The philosophy of marking also affects the types of symbols or statements used on report cards. A variety of symbols have been adopted by the schools throughout the country. What symbols will be best for a particular school system depends upon: (a) the marking philosophy of that system, that is, comparing children with themselves or with their classmates or both, and (b) the kinds of symbols the faculty believes will give parents a true reflection of children's progress.

Research and experimentation have revealed limitations in commonly used techniques for evaluating and reporting pupil progress. Although research and experimentation have indicated pitfalls, they have also provided alternatives to uncritical passive acceptance of existing programs of reporting that accomplish less than what is desired. Those alternatives are for teachers and research workers to appraise continuously evaluation programs and to persist in their attempts to determine the effectiveness of methods for reporting pupil progress.

Probably the best plan of reporting is a combination of the different methods geared to the needs of the local school system. Parents can be active participants in the evaluation process. Cooperative participation is the key concept in any program for improving marking and reporting. Educators, parents, and pupils need to first study the school's educational objectives and then determine the most effective way to communicate the results of pupil progress in terms of them. Only after thorough study, trials of tentative suggestions, critical reactions, and revisions can a school hope to effect a defensible system of marking and reporting.

RESEARCH ON MARKING AND REPORTING

The report card is a major means by which the school reports pupil progress to parents. Teacher marking and reporting systems are likely to continue to fill this role in one form or another. Apparently, many educators believe, as does Rothney, that teacher marks, although often unreliable and invalid indexes of growth, are indispensable tools (95).

Educators appear to be more concerned than parents about the inadequacies of marking and reporting systems. Educators are apparently more aware of the limitations of typical grading schemes for evaluating progress toward complex educational objectives. In contrast, many parents seem content with the traditional report cards and letter grades and even desire them.

A study by Morris, for example, revealed that the majority of parents of fifth- and sixth-grade pupils preferred letter grades and disapproved of statement type reports (77). Yauch reported that parents tend to prefer whatever type of reporting system is currently employed (122). Richardson points out that parents often have unwarranted confidence in the precision with which grades in school can indicate a child's ability or foretell his success in adult life (91).

Although many difficulties continue to plague the teacher and the administrator in evaluating and reporting pupil progress, some real progress has been made in improving reporting practices. It is now common, for example, to provide separate scales for judging academic achievement and for dealing with attitudes, personl-social adjustment, and other noncognitive aspects of development. Again, many reporting systems make provision for conveying to parents other significant information--information about levels of reading achievement, health practices, and the like (60).

In recent years, the dual marking system appears to have gained a number of supporters. Such systems provide two grades: one reflecting the child's achievement in terms of the group norm, and the other, his achievement in terms of his own potentialities and abilities (9, 46, 94). An advantage of such systems, presumably, is that they make it easier for a parent to understand the meaning of the child's grades (45).

Another trend is increased awareness among school personnel that regardless of how complex the reporting form may be, it must be supplemented by other contacts with the home. Telephone conversations, various types of correspondence, parent-teacher interviews, and collections of a pupil's work are commonly employed means by which school and home exchange pertinent information about children. Many writers have emphasized the importance of such procedures (16, 29, 40, 59, 72, 88, 117).

A number of writers suggest that it is a wise policy for the school to solicit the help of parents in devising marking and reporting systems (88, 122). It seems likely that such a practice might acrue benefits in the area of improved public relations.

Marking systems are often viewed as a system of communication. According to Ebel, only to the degree that the marks do have the same meaning for all who use them is it possible for them to serve the purposes of communication (33). That teachers vary considerably in their marking practices has been widely discussed.

There is a paucity of research on procedures designed to create greater uniformity. It has been suggested that prospective teachers be required to take courses in educational measurement; that in-service training programs will help solve the problem; that joint committees of teachers, supervisors, and administrators could help to develop criteria and guidelines. There even seems to be a segment of professional educators who would do away with grades entirely.

A further complication in making reports meaningful results from increased use of ability grouping. Schools which use special "tracks" or cross-class grouping procedures often have difficulty assigning a letter or number grade which is at the same time realistic and meaningful. Few adequate models have been developed and reported in the literature.

Fundamentally, the difficulty associated with marking in an ability grouping situation seems to involve which base group should be employed in comparing the pupil's achievement. Should the student in the top group be compared only with those in his own group or with all pupils in his grade level? Similarly, should the child in the lowest group be judged according to his classmates or in relation to the achievement of all pupils in the same grade? In attempting to deal with such problems, many

schools resort to multiple grading systems which allow the teacher to utilize all of the letter grades of the system but to specify the group on which the grade is based. In such instances, the multiple grading system is supplemented with narrative statements about the child's level of performance to insure that all concerned understand the meaning of the assigned grade. Doak proposes that grouping of students does little to reduce the variability in achievement as indicated by school marks and indicates that a dual marking system may offer a solution to the problems of marking when "track" systems or ability grouping procedures are employed (31).

Teacher marks and report cards furnish one means by which learning is reinforced. They should not, however, constitute the sole or even the major means or reinforcement, nor should grades alone be made major incentives. Little or no experimental work concerning this problem is reported in the literature. Presumably, the lack of discussion implies a belief among educators that other techniques are more valuable in stimulating and motivating pupils--the evaluation of daily assignments, the use of praise or reproof, continual pupil-teacher conferences, and the like.

It may be, however, that the impact of grading practices and reporting systems on pupils has been somewhat taken for granted. If human learning is largely goal directed, it seems reasonable to ask the learner to help in establishing realistic goals and to participate in evaluating his progress toward these goals. In this conviction, Pemberton suggests that pupil motivation can be strengthened by having students participate more actively in marking. He believes that a pupil should rate himself by checking his performance on an inventory which includes such self-evaluations as achievement on tests and assignments as well as habits of study and plans for self-improvement (83).

Boyd points out that children are not necessarily motivated by marks (16), and Alexander believes that it is necessary to determine whether or not a marking system enhances the pupil's self-concept (4).

A number of writers have reported that girls tend to be graded more favorably than boys for similar levels of achievement (19, 48, 68, 97). The effect of such practices on motivation of boys can only be assumed.

Finally, Rolf notes that pupil progress reports, regardless of form, are not going to solve the problems associated with achievement or lack of achievement (94).

Holland points out another weakness in any teacher rating system (55). In a study which attempted to identify the factors which influence teachers to rate pupils favorably, he found that students rated high by their teachers appear to be bright, persistent, conscientious achievers and pupil leaders. He cites further the operation of the halo effects which become apparent in rating and marking.

Ludeman reports that many students are given higher marks than they deserve because they conform to the teacher's expectations in social rather than academic areas (71). Both teachers and pupils confuse their evaluative judgments with personal needs to justify, excel, defend, enhance, punish, or depreciate themselves. These personal needs, according to Symonds, destroy the completely objective significance of evaluation (111). Where letter grades are employed, there is great variablity among teachers in the same school and among teachers within the same system. The way the teacher perceives the child affects his assignment of marks (118).

In determing what grading and reporting system is most satisfactory, the findings of research may prove useful. Here are the results of the range of various studies:

- As a summary of a child's progress in school, some kinds of report cards fall short of the mark, while others give the parent useful information, according to a study by Chansky (20). In this study 70 elementary-school teachers rated 25 reporting devices selected at randon from those used in New York elementary schools. The commonly accepted goals for elementary education were used as criteria, and cards were judged on an eight-point scale. Three main features distinguished high- from low-rated cards. First, the number of items on the card was a significant factor; for example, the more student characteristics to be evaluated on a card, the higher was its rating. Another important factor was the range of behavior covered by the card. The highly rated cards dealt not only with subject matter but also with social adjustment, work habits, and health. Finally, the system of reporting used was significant; the highest-rated cards used letters or numbers with descriptive sentences so that teachers were stimulated to describe the child's achievement.

- Although some educators have criticized the use of grades at all levels, a recent study does not confirm this opinion (52). Grades are a spur to classroom performance, according to Hawk and DeRidder, who compared the performance of pregraded students with that of grade-motivated students. A group of students were awarded grades in a course at the beginning of the term on the basis of previous grade point averages and compared with a control group whose grades were determined at the end of the term on the basis of test scores and a term

project. The removal of pressure for grades acted as a depressant on the work output and test scores of the experimental group. The control group was superior in performance and the researchers concluded that in the absence of motivation for a grade, the majority of the experimental group did not exert their usual degree of effort.

● Some studies have been conducted to try to determine whether students graded on an individualized basis were motivated to do better work than those graded by the traditional method (10). On the whole the results showed that there were no significant differences in achievement among students graded by the different methods. Researchers suggest that there may be two possible explanations for these results. One is that motivation is not significantly affected by methods of reporting. The other is that the reporting of pupil progress was not essentially different in the different programs despite the stated differences in philosophy. In other words, the basic assumption behind the philosophy of individualized reporting is the ability of the teacher to evaluate student effort. This is often called subjective grade. The extent to which the teacher feels the student is or is not working up to his capacity as measured by an IQ test or teacher estimate is usually the basis for the grade. A recent study by Halliwell also tested whether teachers in a school system that used the individualized reporting method really adhered to such a philosophy (46). Results showed that there was a significant relationship between IQ and grades. The conclusion was that teachers tended to mark students in the traditional way in spite of the stated reporting philosophy.

● Halliwell and Robitaille have tested the relation between theory and practice in a dual reporting program (47). In theory, the researchers stated, the dual reporting system seems ideal, but this study showed a wide diversity between theory and practice. With this form of reporting there should be a difference between the grade the student receives in terms of grade norms and the subjective grade he receives in terms of his own ability. But results indicated that teachers tended to give good grades to bright students on both parts of the two-way report. The extent of this halo effect showed that the values that have been attributed to the dual report are hard for teachers to achieve. In effect, bright pupils are being rewarded with good grades twice and slow pupils are being penalized with poor grades twice. The significantly high relation between IQ and individual grades clearly indicated that the teachers in this study were unable to appraise effort in an adequate way. Whether in-service training and orientation in dual reporting for teachers will solve the problem or whether more accurate evaluation techniques must be devised is still open to question.

● Alpren recommends the dual marking system, one grade for achievement based on the student's performance according to general standards and another that would emphasize effort or achievement based on the pupil's potential ability to perform up to generally accepted standards (5). He says this method was used some years ago but, at the time, the first grade was designated by an alphabetical letter and the second was usually expressed as a number; he believes this lessened the importance of the second grade. Alpren recommends that similar symbols (alphabetical letters) be used for both grades because this allows for more objective reporting by teachers. He also recommends that teachers be made aware of the reasons for having both grades in the same form and that the new grading policy be explained to lay and pupil populations through meetings and news releases. The meaning and results of the plan should also be made known to colleges.

● In discussing grouping practices in junior high school, Baughman and Schoonmaker describe several methods schools were using to appraise and report pupil progress under grouping (11). Some educators believe that the slow learner should not be discouraged by low marks, and there are others who think the gifted pupil should not be penalized by lower grades that come as a result of competition in the high-level group. Several suggestions have been made, and schools have solved the problem in various ways. One suggestion was to keep the regular grading system; that is, pupils in talented groups would get A's and B's, the average group would get B's and C's, and the low achievers would get C's and D's. Other suggestions and methods made use of slightly different forms of the dual marking system. One school gives the full range of grades at each ability level but uses sub-numbers--1, 2, and 3--to identify the high, middle, and low levels in each group. Some schools use a system similar to this one, but the numbers range from 1 to 5 and show the effort of the student. The meaning of each number is classified by a descriptive statement. A junior high school uses a double-grade plan with letter grades for both. One grade is given for comparative achievement based on broad standards and the other for personal achievement. The latter grade indicates what each pupil has been able to achieve with the material given to him, regardless of the grade level of the material. The honor roll is based on the pupils' personal achievement grades. Another school system used what is called a job-analysis plan that was initiated after a 10-year study. Every educational task from grades 1 through 12 was analyzed and goals were set up. These goals were listed on the report card, and pupils were checked as to capability in all specified areas each grading period. This check system was accompanied by a letter grade.

● A philosophy of reporting was developed at Alamosa (Colo.) High School that was based on a recognition of individual differences instead of set standards (110). Each student was

judged on his own growth as well as his ability to meet standards. The dual report was only a part of the entire system. Interesting features were that marks accumulated on cards, and parents saw all previous marks recorded rather than those for just one reporting period; and that each student took his card from class to class and personally received grades in conferences with each teacher. Results were favorable; the number of honor roll students doubled, absences of students greatly decreased, and discipline problems were fewer.

Current Practice

About three teachers in four use a report card with a classified scale of letters for reporting pupil progress to parents. About two teachers in five use a scheduled conference. (The data reflect the fact that a teacher may use more than one reporting method.)

These are findings of a nationwide sample survey of public-school teachers conducted by the NEA Research Division in spring 1969 (80). In answer to the question, "What method do you use to report pupil progress to parents? (Check ALL that apply.)," the respondents replied:

A report card with a classified scale of letters (e.g., A-F) ...	71.6%
A scheduled conference with parents	40.2
A written description of the pupil's performance	17.4
A report card with a classified scale of numbers (e.g., 1-5) ...	9.4
A report card showing percentage grades	6.2
A report with either pass or fail.	5.4

Analysis of responses revealed differences in the distribution for elementary- and secondary-school teachers. A higher percentage of elementary than of secondary teachers use a report with either pass or fail, a written description of the pupil's performance, and scheduled conferences with parents. The percentage using report cards with classified letter scales and report cards with percentage grades is higher among secondary than among elementary teachers.

	Elementary teachers	Secondary teachers
A report card with a classified scale of letters	71.6%	83.1%
A scheduled conference with the parents	59.9	20.0
A written description of the pupil's performance	24.3	10.4
A report card with a classified scale of numbers	10.0	8.8
A report card showing percentage grades	2.4	10.0
A report with either pass or fail	8.2	2.6

Reporting methods of teachers in different size school systems differed on two points. The largest school systems (25,000 or more pupils) had a higher proportion of teachers (80.2 percent) using report cards with a classified scale of letters, as compared with a smaller proportion (75.7 percent) in medium-size systems (3,000-24,999 pupils). School systems with fewer than 3,000 pupils showed a higher percentage (43.3 percent) of teachers using a scheduled conference with parents than systems with 25,000 or more pupils (36.7 percent).

	25,000 or more pupils	3,000-24,999 pupils	Fewer than 3,000 pupils
A report card with a classified scale of letters	80.2%	75.7%	77.1%
A scheduled conference with parents	36.7	40.2	43.3
A written description of the pupil's performance	17.6	18.8	15.1
A report card with a classified scale of numbers	10.6	10.0	7.4
A report card showing percentage grades ..	5.6	5.5	7.6
A report with either pass or fail	4.9	5.8	5.3

Geographic analysis of responses showed major differences. A much smaller proportion of teachers in the Southeast (26.7 percent) than in any other section of the country (Northeast, 42.6 percent; Middle, 44.0 percent; and West, 43.7 percent) used scheduled conferences with parents. A higher proportion of teachers in the Southeast (89.3 percent) than in any other section of the country (Northeast, 66.7 percent; Middle, 81.0 percent; and West, 74.7 percent) used a report card with a classified scale of letters. A report card showing percentage grades was used by a higher proportion of teachers in the Northeast (16.0 percent) than in any other section of the country (Southeast, 3.5 percent; Middle, 1.2 percent; and West, 3.4 percent).

In November 1967, the NEA Research Division sent a questionnaire to a sample of 1,103 public school systems having kindergarten (79). Seven basic types of reports were listed and respondents were asked to check those used in their school systems. No effort was made to

TABLE 1.--TYPICAL METHODS FOR REPORTING PUPIL PROGRESS TO PARENTS, 1967-68

Reporting methods	Enrollment grouping				Weighted estimates, all systems operating kindergarten, 1967-68
	Group A 25,000 or more	Group B 3,000-24,999	Group C 300-2,999	Group D 1-299	
1	2	3	4	5	6
Conferences	90.8%	92.3%	91.1%	71.3%	85.6%
Letters to parents	33.0	37.1	38.3	23.3	33.7
Checklist report	35.8	37.4	36.1	22.9	32.5
By telephone	33.0	29.4	24.6	11.2	21.7
Written report using letter grades	16.5	10.5	11.2	24.7	15.0
Written report using percentage marks	...	...	...	0.4	0.1
Written report using descriptive words	33.0	40.3	45.0	39.0	42.4
Other	5.5	1.9	1.0	2.2	1.5
Not indicated	2.8	...	0.6	0.9	0.6
Total number of systems	109	313	313	223	9,766

Source:
National Education Association, Research Division. Kindergarten Education in Public Schools, 1967-68. Research Report 1969-R6. Washington, D. C.: the Association, 1969. p. 30, Table 17.

secure qualitative evaluation, only frequency of use.

Table 1 shows a frequency distribution of the methods typically used to report pupil progress to parents of kindergarten children, the most common type being the scheduled conference.

Public school systems tended to use a combination of methods for reporting the progress of kindergarten children to their parents. The combination consistently included the parent-teacher conference. For the total systems, the most common types of combination reports to parents were, in order of their use:

- Conference plus descriptive word report
- Conference plus informal letter
- Conference plus written check list

Recent Research, 1960 to the Present

This section abstracts 29 selected research studies on the effects of grading and reporting which have been published since 1960. The results of these studies are presented in somewhat general terms. The original study should be consulted for specific conclusions.

It is advisable to inform the reader of some of the criteria applied in selecting the studies to be included. The reader should also be cautioned about potential restrictions on the use of these abstracts of research studies.

First, studies concerned with practices from grade 1 through graduate school were considered. Any studies published prior to the period beginning with 1960 are not included. Furthermore, only sources generally accessible to teachers and administrators were consulted. And finally, an effort was made to select studies with a variety of research emphases.

Each abstract presented here is cast in a standard format: title, purpose of the study, method and procedure, conclusions, remarks, and source. Note that the remarks section represents the point of view of the author of the particular research study under consideration.

There are a number of restrictions on the use of these abstracts. First, although the studies are drawn from the period inclusive of 1969, very recent research is not included since the results of many recent studies have not yet been published.

Second, the reader should be wary of attempting to make facile comparisons of the results of studies for the purpose of making a set of definite generalizations. The diversity of the abstracted studies and the problems of equating and synthesizing research findings in grading and reporting make such an attempt futile.

Finally, these abstracts of research related to grading and reporting could be extremely useful in developing organizational change. However, imitations of successful programs or modification of existing programs in order to conform with or depart from research findings can be made only after careful study of the individual circumstances of the local school situation.

The 29 abstracts are presented in two sections. The first section (beginning on this page) deals with 10 research studies related in grading and reporting in the elementary and secondary school. Studies of grading and reporting at the college level are presented in the second section (beginning on page 37). All studies in each section have been arranged in alphabetical sequence by author. Table 2 categorizes the 29 selected recent research studies with regard to the effects of marking and reporting on eight specific areas.

STUDIES OF MARKING AND REPORTING IN THE ELEMENTARY AND SECONDARY SCHOOL

Elementary-School Teachers Rate Report Cards

Purpose: The purpose of the study was to determine whether teachers believe that various reporting forms are equally effective in relaying functional information about the growth of elementary-school children.

Method and Procedure: Twenty-five elementary school reporting devices were selected at

TABLE 2.--CHART OF SELECTED RESEARCH ON THE EFFECTS OF MARKING AND REPORTING

Research study	As measured by behavior and/or tests in	
	Elementary and secondary school	College
1	2	3
1. Preference for or rating of report forms	Chansky (1963) Chansky (1965) Gillcrist (1965)	Bowers (1967) Sgan (1969)
2. Retention or failure	Christensen (1968) Kamii (1963)	Philbrick (1968)
3. Teachers' subjective grade evaluations	Miner (1967) Storey (1968)	Aiken (1963) Harnett and Stewart (1965)
4. Pupils' perceptions	Flanders (1968)	Bostrom (1961) Burke (1968) Karlin and others (1969)
5. Relationship of measurement devices and reporting forms	Halliwell (1960) Halliwell (1962)	Stroup (1966)
6. Competition for grades		Clark (1969) Hawk and DeRidder (1963) Karlin and others (1969) Mannello (1964)
7. Students' predictions of their grades and student participation in grading		Burke (1969) Garvin (1967) Keefer (1969) Murstein (1965) Robertson (1960)
8. Grades as predictors of success		Michall and others (1962) Stricker (1967)

NOTE: All the studies above are listed in the references at the end of the Research Summary.

random. Seventy teachers from more than 30 school systems were asked to evaluate these devices. Each reporting device was to be rated according to the following goals for elementary school: (a) organismic development; (b) social development; (c) geopolitical development; (d) development in awareness of natural environment; (e) aesthetic development; (f) development of verbal communication; (g) development of nonverbal communication. Each reporting device was judged on a scale of eight points.

Conclusions: The ratings tended to cluster about a point on a continuum rather than being distributed in a spurious manner. An examination of the report cards revealed at least three features that distinguished high from low rated cards: (a) the number of entries; (b) the range of child behavior covered by the card; and (c) the system used in informing parents of a child's progress in school.

Remarks: The study was conducted in New York City.

Source: Chansky, Norman H. "Elementary School Teachers Rate Report Cards." Journal of Educational Research 56: 523-28; July-August 1963.

The Effect of Discontinued Grade Reporting on Pupil Learning

Purpose: This study was designed to test the hypothesis that if grades and the possibility of failure are removed, there will be little impetus for students to achieve prescribed learnings.

Method and Procedure: The sample for this study consisted of 24 pupils in eighth-grade mathematics. During the first semester, each pupil received a letter grade. At the onset of the second semester, the pupils were informed that each would be given a grade of P (pass) as the permanent grade for the second semester. Throughout the second semester there was no modification of teaching techniques or devices.

Conclusions: A high positive correlation was obtained between the rankings of the first and second semesters. Standardized measurements showed that during the second semester, pupils learned more than a typical pupil during an entire 10-month school year. The consistency of these standards to those of teacher rating further supported a high positive correlation between equivalent scores at the beginning and close of the second semester. Observations by teachers indicated that the pupils maintained keen interest in mathematics during the second semester. Pupil reaction, obtained through an opinionnaire, reflected support for the traditional letter grade system used during the first semester. Because of the small size of the sample and the homogeneous nature of the pupils, the author cautioned against drawing definite conclusions from his study.

Remarks: This study was conducted at the Campus Laboratory School at Wisconsin State University in Eau Claire.

Source: Christensen, Donald J. "The Effect of Discontinued Grade Reporting on Pupil Learning." Arithmetic Teacher 15: 724-26; December 1968.

Preferred Items on Pupils' Report Cards

Purpose: The primary questions of this study were to determine what particular items adults prefer to see on pupils' report cards and how parents' and teachers' preferences might differ.

Method and Procedure: An inventory of report cards was administered to 71 parents and teachers. The inventory directed the subjects to comment on their own children's school progress; to indicate their preference for combined or separate judgments for knowledge, skills, and attitudes; to check the number of items they would like to see on a report card; and to check the kinds of items they would like to have a report card contain. The check list contained nine curricular areas common to the elementary school.

Conclusion: Although the subjects stated they preferred cards with 10-15 items, they selected at least 22 behaviors for evaluations. Among the areas emphasized were reading, arithmetic, conduct, and social-emotional development. Knowledge items were more popular than were skill and attitude items.

The differences between parent and teacher item preference were negligible. All parents preferred to have teachers inform them about their children's knowledge rather than skills or subject-matter ideals.

Remarks: The inventory of report cards was administered at a PTA meeting in Greenville, N. C. The occupation of the principal wage earners in 55 percent of the subjects' families was "business," and in 45 percent "professional."

Source: Chansky, Norman M. "Preferred Items on Pupils' Report Cards." Education 86: 169-73; November 1965.

Three Factors of School Achievement

Purpose: This study examines teachers' subjective grade evaluation in relation to a series of more objective achievement assessments including standardized achievement tests and

intelligence tests in an attempt to determine the underlying structure of such evaluations.

Methods and Procedures: Assessments of achievement at different points in a pupil's academic career were obtained for 671 students in three high schools in a midwestern city. The structure of achievement assessments was studied empirically by subjecting 20 different achievement measures to a factor analysis. Included among these measures of achievement were teachers' achievement test scores and intelligence test scores. Each measure was obtained for all subjects at different times in the pupil's academic career. If achievement were a unidimensional concept, all of the achievement measures should define a single factor.

Conclusions: Three clearly defined factors emerged from the analysis of 20 achievement variables: (a) objective achievement defined by intelligence measures and standardized achievement test scores; (b) early citizenship defined by early measures of citizenship and marks; and (c) high-school achievement defined primarily by ninth- and twelfth- grade marks. The high-school achievement cluster showed some relationship to the objective achievement cluster. Citizenship and early marks had some dependency on objective achievement, but appeared also as a distinct factor. The author stated that the most surprising thing in this structure was the absence of more dependency among the factors.

The content of early citizenship suggested that the teachers' evaluation in the early grades tended to be assessments of behavior rather than academic performance. The relationship between early school marks and later high-school marks was relatively low. The joint loadings of these variables were small. Although each of the subjective achievement clusters was related to the objective achievement, these loadings were also small. The author concluded, therefore, that children are assessed relatively independently on the three aforementioned factors as they proceed through school.

Remarks: The sample pupils for this study were enrolled in high schools in a small midwestern city containing a large university. The only criterion for inclusion in the sample was that pupils had begun their public-school education in a city public school in 1951 and had graduated with their class in 1964.

Source: Miner, Betty Crowther. "Three Factors of School Achievement." Journal of Educational Research 60: 370-76; April 1967.

Changes in Pupil Attitudes During the School Year

Purpose: This study was designed to test the hypothesis that the perceptions of pupils toward their teacher and their class activities change during the school year.

Method and Procedure: A Michigan Student Questionnaire (MSQ) was administered to 101 sixth-grade classes in 15 school districts. Thirty classes were selected for further study to include the top 10, the bottom 10, and the 10 near-average for the 101 classes. The MSQ was readministered in January and again in May in these 30 classes (N=800 pupils). Possible change in pupil attitude was related to the following variables: pupils' IQ, socioeconomic status, percentage of A and B letter grades, "externality" or "internality" of the pupils, and the teachers' verbal classroom behavior.

Conclusion: The mean percentage of A and B letter grades for the high-change group was 56.5 and for the low-change group, 64.5. A "z" test between independent proportions yielded a value of 1.66, which was not significant at the .05 level. These data suggest that changes in class attitudes are not significantly associated with grades given by the teacher.

An erosion of positive attitudes was not related to pupils' IQ, socioeconomic status, or percentage of A and B letter grades, but was related to the "externality" or "internality" of the pupils' and to the teachers' verbal classroom behavior. Greater losses in attitudes occurred among external than among internal pupils and among pupils whose teachers exhibited a lower incidence of praise and encouragement than among those whose teachers exhibited a higher incidence of such behaviors.

Remarks: Data from two separate studies were compared. One study was conducted in Minnesota in 1960-61, and the other was conducted in 1964-65 with the MSQ technique.

Source: Flanders, Ned A.; Morrison, Betty M.; and Brode, E. Leland. "Changes in Pupil Attitudes During the School Year." Journal of Educational Psychology 50: 334-38; October 1968.

A Study of Grading in Arithmetic

Purpose: The purpose of this study was to determine whether fourth- and fifth-grade teachers preferred an S - N marking system in arithmetic over the traditional system.

Method and Procedure: A new arithmetic text and program were introduced in grades 4 and 5. At the same time, teachers were directed to mark students work S for satisfactory and N for needs additional help. Teachers were then queried concerning their reactions to the S - N marking system.

Conclusions: The majority of teachers (31 vs. 26) favored use of the S - N marking system.

These teachers were the more experienced (10.6 vs. 5.5 years). The author suggested that the philosophy and procedures for evaluating all students' work need to be re-examined.

Remarks: This study was conducted in Middletown, New Jersey.

Source: Gillcrist, William A., Jr. "A Study of Grading in Arithmetic." Education 86: 177-81; November 1965.

Dangers Inherent in Correlating Averages

Purpose: The purpose of this study was to corroborate the hypothesis that correlation with averages is questionable in the area of teacher grading, particularly in the practice of marking on an "individualized" basis.

Method and Procedures: The sample was limited to fourth- and sixth-grade pupils--352 boys and 313 girls in 29 classes in the fourth grade, and 300 boys and 290 girls in 21 classes in the sixth grade. An intelligence test was administered to ascertain the IQ's of the students. The report card of each pupil was analyzed to determine whether the pupil had received satisfactory or unsatisfactory grades. Separate tabulations were made for boys and girls.

Conclusions: The numbers and percentages of boys and girls receiving satisfactory and unsatisfactory grades according to IQ classifications at the fourth and sixth grade levels showed a positive correlation between grades and IQ for both sexes at both grade levels. A determination of biserial coefficients of correlation indicated that the correlations were all signifcant at the .01 level of confidence. Product-moment coefficients of correlation between the average IQ's and the average passing rates in the classes resulted in negative correlations at the fourth-grade level and positive correlations at the sixth-grade level. The author concluded that the results of this study tend to support the hypothesis that correlation with averages is particularly questionable in the area of teacher grading.

Remarks: This study was conducted in the public elementary schools of a suburban Long Island school district.

Source: Halliwell, Joseph W. "Dangers Inherent in Correlating Averages." Journal of Educational Research 55: 327-29; April 1962.

The Relationship of Certain Factors to Marking Practices in Individualized Reporting Programs

Purpose: The purpose of this study was to determine the relationship of IQ, academic achievement, and learning efficiency to report card grades in a district which had adopted an "individualized" type of marking program in the elementary schools.

Method and Procedure: The school district in which this study was undertaken had adopted the "individualized" system (evaluating pupil progress on the basis of the pupil's ability rather than grading on the basis of class or national norms) of reporting pupil progress in the elementary schools. The subjects of this study were all the fourth- and sixth-grade pupils in the district with the exception of those who were absent when an intelligence test was administered, those who were absent when an achievement battery was administered, and those who entered the term late. The subjects were 665 pupils in 29 classes in the fourth grade and 590 pupils in 21 classes in the sixth grade.

The Pintner Intermediate Test, Form A (Verbal), and the Stanford Achievement Test, Intermediate Battery-complete, Form D, were administered to all subjects. A deviation IQ and a median grade equivalent were calculated for each pupil. The pupils' report cards were analyzed to determine who had received satisfactory marks and who had received unsatisfactory marks. Coefficients of correlation between IQ and median grade equivalents in achievement were calculated at both the fourth- and sixth-grade levels. Regression equations of achievement on intelligence were then set up for the fourth and sixth grades, and an effort quotient was determined to serve as the measure of learning efficiency. Separate biserial coefficients of correlation were calculated between intelligence and satisfactory or unsatisfactory grades, achievement and satisfactory and unsatisfactory grades, and learning efficiency and satisfactory and unsatisfactory grades at the fourth- and sixth-grade levels for boys and girls.

Conclusions: Based on the findings of this study, the author believed that the following conclusions seemed warranted:

a. Intelligence was significantly related to grading practices at the 0.1 level of confidence. This correlation between intelligence and grading closely approximates the usual findings concerning the relationships between intelligence and school marks under a traditional grading system.

b. The correlations between median grade achievement scores and grading practices were slightly higher than were those between intelligence and grades. This finding closely parallels the results of studies employing traditional grading systems.

c. No relationship existed between learning efficiency and grading practices.

d. Although the teachers employed in the school system had ostensibly adopted an "individualized" marking program, the significant positive relationships between intelligence and grading, and standardized achievement scores and grading lead to the conclusion that marks were assigned in a traditional manner irrespective of the stated marking philosophy of the school district.

e. Any investigations to determine the superior motivational effectiveness of different types of marking procedures must first determine if the teachers involved are adhering to the stated marking philosophy of the school district.

Remarks: The study was conducted in a public school system in suburban Long Island, New York.

Source: Halliwell, Joseph W. "The Relationships of Certain Factors in Marking Practices in Individualized Reporting Programs." Journal of Educational Research 54: 76-78; October 1960.

Marks, Achievement, and Intelligence of Seventh-Graders Who Were Retained (Nonpromoted) Once in Elementary School

Purpose: The purpose of this study was to determine if the retained elementary-school pupils will come up to the level of the promoted pupils.

Method and Procedure: The retained group was defined as the pupils who had been in the school system since kindergarten and who had a record of one (but not more than one) retention in grades 1 through 5. The control group was defined as those who had been in the school system since kindergarten and who had never been retained and were in the expected age grade. Each group had 22 boys and 9 girls. Three variables were studied: (a) marks received at the end of the first semester in seventh grade in five academic subjects, (b) achievement test scores in reading and arithmetic, and (c) intelligenge test scores.

Conclusions: The retained pupils were found to have the following characteristics: (a) Their marks in academic subjects were significantly below the average of the promoted pupils, and the majority of their marks were D's and F's. (b) Their achievement levels in reading and arithmetic were significantly lower than those of the promoted pupils. (c) Their IQ's were significantly lower than those of the promoted pupils, but more than half of the retained pupils had at least average IQ's (as measured by the California Test of Mental Maturity). (d) The reason for getting D's and F's can be attributed neither to low intelligence nor to poor basic skills.

Remarks: This study was conducted during the 1959-60 academic year in the two junior high schools in Ypsilanti, Michigan.

Source: Kamii, Constance K., and Weikart, David P. "Marks, Achievement, and Intelligence of Seventh Graders Who Were Retained (Nonpromoted) Once in Elementary School." Journal of Educational Research 56: 452-59; May-June 1963.

Some Evidence in the Essay-Item Case

Purpose: This study was conducted to test the hypothesis that in the absence of any real ability to judge the worth of a written response, teachers grade written responses against an internalized "set" based on a symmetrical distribution around a given mean.

Methods and Procedures: Three paragraphs were developed and judged to be excellent, fair, and very poor by two criterion groups using different approaches. Certified, experienced teachers (N=261) were then asked to award percentage grades to the paragraphs.

Conclusions: There was no statistical significance in the difference in means, standard deviations, or shapes of the distributions of grades obtained from the three paragraphs. The author concluded, therefore, that the grades awarded for essay item response reflected a teacher set rather than the value (inherent) in the examinee's response.

Remarks: The teachers rating the paragraphs were not permitted to communicate with each other during the marking period.

Source: Storey, Arthur G. "Some Evidence in the Essay-Item Case." Journal of Educational Research 61: 351-54; April 1968.

STUDIES OF MARKING AND REPORTING AT COLLEGE LEVEL

The Grading Behavior of a College Faculty

Purpose: The purpose of this study was to show that teachers grade with reference to the existing ability level of their students.

Method and Procedure: In 1959, the college began selecting freshman students on the basis of a multiple regression equation, which consisted of assigning numerical weights (based on the 1958 freshman class) to three predictor variables: Scholastic Aptitude Tests--Verbal (SAT-V), Scholastic Aptitude Tests--Mathematical (SAT-M), and a converted two-digit score of rank in high-school graduating class (HSR). The equation, Predicted Grade (PG) = .037 (SAT-V) + .010 (SAT-M) + .328 HSR - 21.98, yielded a multiple correlation of .70 and was used to predict

freshman grade averages. Student scores derived from this equation were studied and compared for the years 1959, 1960, and 1961.

Conclusions: For the years 1959, 1960, and 1961, the mean scores on SAT-V, SAT-M, and HSR for freshman students who were admitted progressively increased, but increases in the means of the predictor variables were not accompanied by an increase in the criterion mean. Although mean SAT-V, SAT-M, and HSR and consequently PG rose steadily from 1959 to 1961, the actual freshman average grade (FAG) mean did not follow suit.

All increases in the means of the predictor variables (SAT-M, SAT-V, HSR, and PG) were statistically significant, by student t-tests, but the criterion mean (FAG) was relatively stable.

The author concluded that the grading behavior of the faculty was not based on standards that held constant over the years. The grading standards shifted with the ability level of the class, being more stringent in each successive year. Thus, the faculty as a whole had no implicit standards of performance. The standard was dictated by the quality of students in the freshman class.

Remarks: This study was conducted at the Woman's College of the University of North Carolina.

Source: Aiken, Lewis R., Jr. "The Grading Behavior of a College Faculty." Educational and Psychological Measurement 23: 319-22; Summer 1963.

Grades as Reinforcing Contingencies and Attitude Change

Purpose: This study was designed to examine the effect on attitudes of differential assignment of grades for performance on attitude-related essays.

Method and Procedure: A 40-item questionnaire containing four 10-item attitude scales was administered to 228 students. During a class period approximately six weeks following administration of the scales, the subjects were asked to write essays on particular assigned topics. Scores on the attitude scales determined the position that was assigned. In each case, the subject was instructed to write supporting the position opposite to that indicated by the pretest scale score. The designation of topic to a particular subject was based on the strength of his initial position on the attitude scales. On a random basis, grades were assigned to the essays. One-third of the subjects writing on each topic received a grade of A, one-third received a grade of D, and one-third was given no grade. Immediately after returning the essays and grades, the attitude questionnaire was readministered.

Conclusions: It was predicted that subjects who were awarded an A would change on the average in the direction of their essays to a greater extent than subjects who were given a D. Subjects receiving an A changed an average of 31.76 points in the direction of their essays while subjects who were given D changed 25.85 points. This difference is significant beyond the .01 level. Comparisons of the groups that received a grade with the group that did not, indicates significantly greater change ($p. < .05$) for the subjects who received an A than for those given no grade, while no difference is suggested between the subjects obtaining a D and subjects receiving no grade.

An analysis of mean change in relation to initial position indicates that those who had initially assumed a favorable position on each of the issues change significantly more ($t = 4.73$, $p. < .01$) than those who were unfavorable.

The results suggest support for the hypothesis that a "good" grade serves to reinforce the behavior for which it had been administered.

Remarks: This study was conducted at the State University of Iowa among students enrolled in communication skills classes.

Source: Bostrom, Robert N., Vlandis, John W., and Rosenbaum, Milton E. "Grades as Reinforcing Contingencies and Attitude Change." Journal of Educational Psychology 52: 112-15; April 1961.

A Test of Variation in Grading Standards

Purpose: This study analyzed the relationships between first-term grades and academic ability of beginning college freshmen in order to determine whether both groups were graded on comparable scales of evaluation.

Method and Procedure: Data were obtained from the records of 4,283 beginning freshmen for the fall of 1963 and 5,132 beginning freshmen for the fall 1964 semester. Two predictors were used (a) high-school percentile rank, and (b) composite score on the American College Test battery. The grade point average achieved in all graded courses completed during the first semester in attendance was defined as the index of grading.

Conclusions: It was assumed that a change in the grading standards for the two groups is reflected in differences in their regression equations predicting first-semester grade point average from a weighted combination of high-school percentile rank and ACT Composite score. On the basis of variance F-tests of common regression and intercepts for the linear regression equations, the author concluded that grading standards had changed from 1963 to 1964.

Remarks: The study was conducted at the Champaign-Urbana Campus of the University of Illinois.

Source: Bowers, John E. "A Test of Variation in Grading Standards." Educational and Psychological Measurement 27: 429-30; Summer 1967.

Some Preliminary Data on the Use of Self-Evaluations and Peer Ratings in Assigning University Course Grades

Purpose: This study reports the preliminary results of a series of investigations designed to explore the effects of increasing the amount of student participation in determining their own grades.

Method and Procedure: Students were selected for this experiment which focused on two methods of course grading: (a) the use of self-evaluations in assigning course grades, and (b) the use of peer ratings in assigning course grades.

Conclusions: Evidence from the self-evaluation method suggested that students were unable to assign their own grades objectively and realistically. There was low agreement between self-evaluations and peer ratings and self-evaluations and instructor ratings. The peer rating method of assigning grades seemed to offer more promise. An indication of the validity of peer ratings was the greater agreement obtained between peer ratings and instructor ratings than between peer ratings and self-evaluations and instructor ratings and self-evaluations.

Remarks: This study was conducted at the University of Minnesota.

Source: Burke, Ronald J. "Some Preliminary Data on the Use of Self-Evaluations and Peer Ratings in Assigning University Course Grades." Journal of Educational Research 62: 444-48; July-August 1969.

Student Reactions to Course Grades

Purpose: The purpose of this study was to consider the issue of grading from the perspective of the student.

Methods and Procedures: The subjects were enrolled in a university course emphasizing application of behavioral science concepts and theories to leadership and management situations. The class met for a 2-1/2 hour period once a week for 10 weeks. The data were obtained from one section of 38 students, 85 percent of whom were seniors. A questionnaire assessing student attitudes toward several issues in grading was administered during the last meeting of the course.

Conclusions: Eight times as many students felt that grading interfered with rather than helped learning (66 percent vs. 8 percent); the remainder (26 percent) were undecided. Students who felt that grading hindered learning believed that the traditional A-F system tended to influence methods of studying and learning, classroom behavior, and the teaching process itself in undesirable ways.

Remarks: This study was conducted in the Department of Management at the University of Minnesota.

Source: Burke, Ronald J. "Student Reactions to Course Grades." Journal of Experimental Education 36: 11-13; Summer 1968.

Competition for Grades and Graduate-Student Performance

Purpose: This study explored the effects of a single concrete source of motivation (competition for grades) on performance of graduate students in education.

Methods and Procedures: The subjects were two intact classes of 49 (Class A) and 59 (Class B) students of advanced educational psychology. One instructor taught both classes, using the same lecture notes, text, assignments, and examination. The students were compared on two criteria: (a) research papers and (b) examination scores. Class A competed for grades on the research papers but not on the examination. The members of Class B were told at the beginning of the course that they would all receive the grade B. Class B was not permitted to compete for grades on either the research papers or the examination.

Conclusion: On all the criteria for the research paper, Class A performed at a significantly higher level than did Class B. The two classes did not differ on their examination scores. Class A, however, was significantly superior to Class B in other, less direct, areas of examination performance (percent taking examination, attendance, hours spent studying, and chapters read).

Within Class A, rough comparisons were made between performances on the research paper and performances on the examination. Under competition conditions, Class A spent an average of 44.5 hours preparing and writing the research paper; under noncompetitive conditions the same students averaged 8.0 hours reading and studying the textbook. Also, in Class A, all of the students turned in a research paper (on which they would be graded) and 84.7 percent took the final examination (on which they would not be graded).

The author believed that these results suggested that performance among graduate students was significantly higher under conditions in

which they competed for grades than under conditions in which there was no competition for grades.

Remarks: The author draws attention to the following possible precautions: (a) sampling bias, (b) experimenter bias, (c) for Class A, a halo effect by which competition on one criterion (research paper) produced competition on the other criterion (examination); and (d) for Class A, the belief that study of the text and class attendance would provide them with a stronger background for completing the research paper.

Source: Clark, D. Cecil. "Competition for Grades and Graduate-Student Performance." Journal of Educational Research 62: 351-54; April 1969.

A Comparison of Students' Predictions of Rank Order on Proximate and Remote Tasks

Purpose: The primary purposes of this study were to investigate the relative magnitude of errors made by high and by low ranking students in their initial predictions of final achievement, to study the effect of relative proximity in time, and to compare the accuracy of predictions made at a remote time with those made at a more proximate time. A secondary purpose was to investigate the extent of disparities between predictions of letter grades and the associated predictions of rank order.

Method and Procedure: The subjects were 63 graduate students enrolled in two sections of the tests and measurement course taught by the experimenter. All subjects were teachers with a median of six years of experience. Within the first five minutes of the first class session, each subject submitted a prediction of the final letter grade (A-F) he would receive in the course, and the final quintile rank (5-highest to 1-lowest) in his own section. The predictions were not examined until the course was over. Midway through the eight-week course, and before any competitive exercises had been undertaken, each subject again submitted predictions of the final letter grade and quintile rank. The day before the examination each subject predicted his quintile rank on the examination, and his quintile rank in a term project. As before, the subjects were assured that their predictions would not be examined until the course was over.

Conclusions: In general, the high third of the group tended to underestimate their prospective achievement a little while the low third overestimated theirs greatly in all predictions.

The major finding of this study was that both high and low ranking students tended to lower their predictions of final ranking by the same amount as the event drew near, and, in predicting their ranks on successive tasks, both groups tended to make much lower predictions regarding the proximate task than the remote task. These two effects were interpreted as a single phenomenon, the general lowering of expectations over time.

An incidental finding was that students' predictions of final letter grades can be ambiguous with respect to their expectation on class standing.

Remarks: This study was conducted at the University of Maryland during the summer of 1966.

Source: Garvin, Alfred D. "A Comparison of Students' Predictions of Rank Order on Proximate and Remote Tasks." Journal of Educational Research 61: 176-78; December 1967.

Personality Rigidity of Students Showing Consistent Discrepancies Between Instructor Grades and Term-End Examination Grades

Purpose: The purpose of this study was to examine the hypothesis that (a) students who generally receive higher grades from their instructors are more conforming, compulsive, and rigid than are students who generally receive higher grades from the term-end examination, and (b) their general academic aptitude scores are lower.

Method and Procedure: Two groups participated in this study: those whose final examination scores were consistently higher than their instructor grades (N = 144), and those whose instructor grades were consistently higher than their examination scores (N = 153). Students were compared on three criteria: (a) actual mean instructor and examination grades, (b) general academic aptitude, and (c) personality rigidity.

Results: The author concluded that in situations where final grades are assigned on the basis of instructor and examination evaluation, students who consistently receive lower examination than instructor grades are more rigid and have less academic aptitude than students receiving higher examination grades.

Remarks: This study was conducted at the University of South Florida in the College of Basic Studies.

Source: Hartnett, Rodney T., and Stewart, Clifford T. "Personality Rigidity of Students Showing Consistent Discrepancies Between Instructor Grades and Term-End Examination Grades." Educational and Psychological Measurement 25: 1111-15; Winter 1965.

A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students

Purpose: The study was designed to determine whether students who had been assigned a final grade for the course at the beginning of the quarter would differ significantly in achievement from students whose grades were to be determined by performance on tests and a term project during the quarter. The null hypothesis was that no significant differences in achievement would be exhibited by the two groups.

Method and Procedure: The subjects were 118 students enrolled in four sections of educational psychology taught by two instructors. One section taught by each instructor was selected as the experimental group, and the other section was the control group. Final grades for the course were assigned to the experimental group (N = 62) early in the quarter, according to previous grade point averages.

Each section was taught by the instructors in an equivalent manner, using the same lectures, discussions procedures, audiovisual aids, and child-study assignments. The students were aware of the group (experimental or control) to which they were assigned. At no time during the quarter, however, did either instructor obtain directly or indirectly any indication which sections had been identified as the control and experimental groups.

Each student's progress was evaluated by a term project, two 60-item examinations, and one 100-item final examination.

Conclusions: Differences between the mean scores for the experimental and the control groups on the two 60-item examinations were significant at the .05 level, while differences for the final examination were significant at the .01 level of confidence.

The term projects were graded on a 10-point scale. The mean number of points for the control group was 6.35, and the mean for the experimental group was 5.81. The difference of .54 was not statistically significant.

Thirty-four members of the experimental groups earned a grade one letter lower than the one assigned; 14, two letters lower; two, three letters lower; and one pupil earned a grade four letters lower than the assigned grade. Grades earned by the control group were well above those earned by the experimental group.

The author concluded that for the experimental group, removal of pressure for a grade acted as a depressant in work output.

Remarks: The study was conducted at the University of Tennessee. The majority of subjects were sophomore and junior education majors, but other colleges were represented with a few students who planned to become teachers.

Source: Hawk, T. L., and DeRidder, L. M. "A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students." Journal of Educational Research 56: 548-50; July-August 1963.

Academic Attitudes and Performance as a Function of Differential Grading Systems: An Evaluation of Princeton's Pass-Fail System

Purpose: The study concerned examining performance in course work and attitudes toward such work as a function of grading systems that differ in their properties (the traditional A-F system vs. a pass-fail system).

Method and Procedure: Students were allowed the option of taking one course per year on a pass-fail basis, while taking all other courses under the traditional grading system. The instructors were not informed which class members were taking courses pass-fail. A questionnaire was distributed to students who had taken pass-fail courses to solict their opinions. This questionnaire was "tested" twice for validity. A second phase of the study consisted of examining academic performance as a function of the differential grading system.

Conclusions: The author concluded that the findings of this study underscored the need for further study and re-examination of the pass-fail system. The need for study and re-examination was reflected in the responses of many students who often believed they learned more, worked closer to capacity, were more motivated to learn and more actively participated in numerically graded courses than in ones marked pass-fail. These beliefs were reflected in actual performance: students received significantly better grades in their competitively graded courses than in their pass-fail subjects. The pass-fail alternative had several advantages: (a) It created a willingness on the part of some students to explore courses they would have ignored if faced with a traditional grade. (b) It reduced tension owing to the elimination of competitive grading. Findings from the study also suggest that a pass-fail system may be undermined when a student's course load creates time pressures in his study schedule.

Remarks: This study was conducted among undergraduate students at Princeton University during the 1966-67 academic year.

Source: Karlins, Marvin; Kaplan, Martin; and Stuart, William. "Academic Attitudes and Performance as a Function of Differential Grading Systems: An Evaluation of Princeton's Pass-Fail System." Journal of Experimental Education 37: 38-50; Spring 1969.

Self-Prediction of Academic Achievement by College Students

Purpose: This report deals with the accuracy and stability of the self-prediction of academic achievement by college students as compared with predictions based on the score of a standardized college entrance test, the high-school GPA (grade point average), and the most recent college GPA.

Method and Procedure: The subjects of this study were 154 liberal arts undergraduate volunteers, who were asked at the beginning of the school year to complete a card on which they listed all of the courses for which they were currently registered and the grade they expected to receive in each course at the end of the first grading period. This procedure was repeated at the beginning of each of the three grading periods of the two-semester school year. Four variables were obtained for each student: (a) the standard composite score on the American College Test, (b) the high-school total GPA, (c) the self-predicted college GPA and (d) the achieved college GPA.

Conclusions: From the data the author concluded: (a) The self-estimate of achievement is just as good a predictor for college students as the pre-college variables in common use. (b) The self-estimate is a significantly better predictor than common pre-college variables when the student has received cues to his achievement, e.g., a midpoint course grade. (c) The self-prediction is a stable variable during a four-year college program. (d) The standardized entrance test and the high-school GPA tend to become less accurate predictors as the student progresses through college. (e) At all points in college the grades most recently earned are the best predictors of those which will next be achieved.

Remarks: The author noted several factors which tended to limit his conclusions: (a) the small size of the sample, (b) the uneven distribution of the subgroups, (c) the effect of uncontrolled variables on achieved GPA, (d) the effect of uncontrolled variables on self-predictions GPA's, and (e) possible effect made by predictions entered on a card.

Source: Keefer, Karl E. "Self-Prediction of Academic Achievement by College Students." Journal of Educational Research 63: 53-56; October 1969.

College Teaching Without Grades

Purpose: The purpose of this study was to determine whether it was possible to substitute non-grading methods of evaluation for grades or scores in a college course without deterioration of the quality or quantity of learning.

Method and Procedure: Students in two education classes (N = 71) were informed that except for the final grade there would be no marks in the course. All tests and papers would be designated either acceptable or unacceptable. Students whose tests were marked unacceptable could change their grade to acceptable by doing a satisfactory assignment on the material unsuccessfully dealt with on the test. Students who got acceptable on all class tests and the final examination would receive a final grade of C. Papers could be presented for consideration of a final grade higher than C.

Conclusions: At the close of the semester, an open-ended questionnaire designed to elicit opinions on the nongrading methods used was completed and returned by 40 of the 71 students. The author concluded that this study demonstrates that the preoccupation of students with extrinsic rewards for study in the form of grades can be reduced. Evidence showed that as a result of the approach used, less cheating took place than in other college courses; students felt less tension in connection with class tests; student opinion changed with regard to their conception of the function of a test; students maintained both the quality and quantity of their academic performance; and finally, ample provision was made for individual student needs and interests.

Remarks: This study was conducted at Hofstra College in the spring of 1962.

Source: Mannello, George. "College Teaching Without Grades." Journal of Higher Education 35: 328-34; June 1964.

High-School Record and College Board Scores as Predictors of Success in a Liberal Arts Program During the Freshman Year of College

Purpose: The purpose of the study was to determine the predictive validity of high-school grade point average, verbal scores, mathematics scores, and total (unweighted) scores of the Scholastic Aptitude Test of the College Entrance Examination Board (CEEB), both individually and collectively, relative to a criterion of grade-point average earned during the freshman year of college.

Method and Procedure: The subjects were 209 freshman men and 223 freshman women who had completed 24 or more units of academic work during their entire first year and no fewer than 11 units during a given semester. A correlation and multiple regression analysis was effected for each sex group.

Conclusions: The study revealed the following: (a) For both sexes high-school GPA is more predictive of success in college than either part scores or total scores of the CEEB. (b) A least squares linear combination of high-school

GPA and CEEB total scores or of high-school GPA and differentially weighted verbal and quantitative CEEB scores yields a higher predictive validity (multiple correlation coefficient) than does any one predictor. (c) The achievement of women in the liberal arts college can be predicted with greater accuracy than that of men.

Remarks: This study was conducted during the 1960-61 academic year at the University of Southern California.

Source: Michael, William B., and others. "High School Record and College Board Scores as Predictors of Success in a Liberal Arts Program During the Freshman Year of College." Educational and Psychological Measurement 22: 399-400; Summer 1962.

The Relationship of Grade Expectations and Grades Believed To Be Deserved to Actual Grades Received

Purpose: The purpose of this study was to explore the inter-relationships between the subject's prediction of grades at the beginning of the term (E 1), his prediction just prior to the final examination (E 2), the grade he believed he "truly" deserved at this time (G d), and the grade assigned to him (G r).

Methods and Procedures: The subjects (Ss) were 76 students comprising four sections of a course in educational psychology. The instructors consisted of a female professor who taught three sections and the author who taught the fourth. There were to be three tests administered during the course. The Ss were asked to predict their final grade at the beginning of the term (E 1), to predict their final grade just prior to the final examination (E 2), and to write down, during the last week of the semester, the grade they thought they deserved (G d) because of their "true" ability. The Ss were assured that the data were for research purposes only, and that participation would in no way influence their final grade. On the basis of the three test scores, a final grade was assigned to each S at the end of the term (G r).

Conclusions: The data were analyzed for the total group and for high (Ss receiving grade A or B) and low (Ss receiving grade $\leq$ C) Ss. The dimensions examined via chi square were change versus no change from one variable to another, and in the case of those Ss who did change, direction of change (Up vs. Down).

The results indicated that high Ss were generally realistic in their expectations and statement of the grade they deserved (the grade received was the criterion). Low Ss tended to be very unrealistic in their initial estimates and to be relatively refractory to the effects of experience. Most low students perceived themselves as deserving a grade of B.

In view of the fact that almost no student perceived himself as a poor or even mediocre student, the author questioned the role of the grade as a positive force in motivation.

Remarks: This study was conducted at Louisiana State University. An experiment to test the motivational effect of giving spuriously high grades was suggested.

Source: Murstein, Bernard I. "The Relationship of Grade Expectations and Grades Believed To Be Deserved to Actual Grades Received." Journal of Experimental Education 33: 357-62; Summer 1965.

Precision in Grading Practices--Panacea or Problem?

Purpose: The study focused on a greater precision in grade-recording (a plus-minus 12-point system) in order to determine any effect upon the retention of those students eligible for academic disqualification.

Methods and Procedures: A random sampling of students, representing each of the academic departments of a college, was selected from the official list of students subject to academic disqualification. Faculty members computed each student's grades on a 12-point plus or minus scale (A+ = 12 to D- = 1). Modified grades were then compared with the original grades of this sample.

Conclusions: It appeared that at least among students on the lower end of the academic achievement continuum, the utilization of a more precise grade-accounting system resulted in reduced GPA's. The author concluded that a more refined system will not tend to facilitate student persistence as was hoped.

Remarks: The study did not include students with high or average grades.

Source: Philbrick, Joseph L., and O'Donnell, Patrick I. "Precision in Grading Practices--Panacea or Problem?" Journal of Educational Research 62: 173-76; December 1968.

Counselor and Student Estimates of Grades as Predictors of Academic Achievement

Purpose: The purpose of this study was to determine how well college freshmen and their counselors can predict first semester grades.

Method and Procedure: Two hundred students and five counselors participated in the study. The data consisted of student predictions of grades before and after testing and counseling, counselor predictions of grades, and the grades actually earned at the end of the semester.

Conclusions: A close correspondence was found between estimates made by students before and after the counseling program. While student past estimates correlated higher with counselor estimates than did student pre-estimates, the correlation of both student estimates with actual grades was approximately the same. Counselor estimates correlated higher with obtained grades than did student estimates. All correlations were statistically significant.

Remarks: The study was conducted at the University of Mississippi.

Source: Robertson, Malcolm H. "Counselor and Student Estimates of Grades as Predictors of Academic Achievement." Journal of Educational Research 54: 73-75; October 1960.

The First Year of Pass-Fail at Brandeis University: A Report

Purpose: The purpose was to make a quantitative and qualitative investigation of the impact of a pass-fail system. Specifically, the study determined how many students, courses, and faculty members were involved, and found also the impression of those who were involved in the system.

Methods and Procedures: Basic data relating to numbers of students, courses, and faculty members were obtained from the faculty and the registrar's offices. A questionnaire was sent to all faculty members who had five or more pass-fail students in any course, and a slightly different questionnaire was sent to a random sample of students who had taken one or more pass-fail courses in the fall or spring term of 1966-67. The pass-fail system was a supplementary part of the evaluative system, and was optional for students of good standing for course work outside their area of concentration.

Conclusions: In the fall and spring of 1966-67, 794 students elected to take pass-fail courses and 98.4 percent of these students received a mark of "pass." Enrollment figures revealed that in the fall term eight courses, or 6.6 percent of those offered, accommodated approximately 32.2 percent of the students choosing pass-fail courses. In the spring term, 10 courses, or 8.4 percent, accommodated 428 percent of those selecting pass-fail courses. Five of the eight and seven of the 10 most popular fall- and spring-term courses, respectively, were in the School of Creative Arts (drama, music, art). The School of Social Science, which had the largest enrollment, had the highest percentage of its concentrators making pass-fail choices; the School of Humanities came second; the School of Science, third; and the School of Creative Arts, last. Students seemed to elect courses in which they had some interest but which had not previously been recognized as possible electives. Respondents to the questionnaire were almost overwhelming in support of continuing the pass-fail system or a closely allied variation of it.

Remarks: Response to the questionnaire was as follows: 71 percent of the sampled faculty members replied, and 48.2 percent of the students who were queried replied to the questionnaire.

Source: Sgan, Mathew R. "The First Year of Pass-Fail at Brandeis University: A Report." Journal of Higher Education 40: 135-44; February 1969.

The Graduate Record Examination and Undergraduate Grades as Predictors of Success in Graduate School

Purpose: This study represents an attempt to assess the relationship between a number of predictors, representing combinations of Graduate Record Examination (GRE) scores and undergraduate performance measures, and a number of criteria of success in a doctoral program in clinical psychology.

Method and Procedure: The subjects (Ss) were 37 students (22 males and 15 females) who had completed all of the course requirements for the doctorate in a clinical psychology program. There were 15 predictor variables, seven derived from GRE scores and eight from undergraduate grades. Four criteria of success in the program were used, three derived from grades and the fourth from the length of time to the completion of the oral examination for the dissertation. All the data were drawn from the permanent records of the Ss. A 19 x 19 correlation matrix was set up to analyze the data.

Conclusions: The single best predictor for all criteria was the grades obtained in undergraduate psychology courses. The most useful GRE score was the sum of the quantitative and psychology tests, although it made only a slight contribution to the prediction of grades, and a negligible contribution to the prediction of the time until orals.

Remarks: This study was conducted at Adelphi University. The Ss represented four consecutive entering classes.

Source: Stricker, George, and Huber, J. T. "The Graduate Record Examination and Undergraduate Grades as Predictors of Success in Graduate School." Journal of Educational Research 60: 466-68; July-August 1967.

Grouping Errors in the Grade-Point Average

Purpose: It was the purpose of this study to make an empirical investigation of the

difference between grade-point average computed on the same transcripts when the computations were made on a five-point scale, C = 2, and a 15-point scale, C = 8.

Method and Procedure: The transcripts of 250 students were selected for analysis. The following selection criteria were applied: (a) a sample of + and - scores, (b) several semesters work for each student, and (c) giving a preference to borderline and probation cases. Greatest attention was focused on the influence of grouping on border line students. To investigate the discrepancies resulting from grouping scores by two methods of computation, it was necessary to use conversion tables.

Conclusions: The author concluded that the differences between GPA's derived by grouping scores into five weighted categories were almost identical with differences based on chance occurrence. The author also stated that neither grouping scale produced consistently higher scores than the other and that the differences produced by one scale could well be related to the manner in which data were collected.

Remarks: The author emphasized GPA's in the critical area of C = 2.0.

Source: Stroup, Francis. "Grouping Errors in the Grade-Point Average." Journal of Experimental Education 34: 31-33; Spring 1966.

Summary

The preceding information points out many weaknesses in current practice and implies some of the steps which should be taken to bring about improvement.

Perhaps the greatest need is for more understanding of the reporting system. The primary purpose of any marking and reporting plan is to communicate to the parent and children the school's assessment of the pupil's progress. Therefore, it is important that the system be understood by everyone concerned--by the parents, the pupils, and the teachers (28). As Hockstad points out, traditional report cards place their entire emphasis on the subject rather than on the learner, and parents have no possible basis for interpreting the marks (54). If a marking system is to be understood, a policy must be formulated which indicates the purpose of the marks, the substance of what is being evaluated, the sources of evidence employed, the basis of comparison, and the curricular reference of the marks (58).

There is also a need to give more attention to the content of reports. Interviews with administrators and teachers reveal that there is more interest in how to report than what to report (116). Even when statements and narrative reports are used to supplement the more formal grade reports, little additional information is conveyed. Teachers often hesitate to make harsh judgments and prefer innocuous statements. It has been suggested that in writing narrative statements, teachers emphasize specific characteristics and behaviors when they describe skills, attitudes, and habits and try to avoid broad, ambiguous statements which say little or nothing.

The literature on reporting to parents suggests several advantageous marking procedures:

1. Assessments of attitudes, conduct, and citizenship need not be emphasized as part of the evaluative marks in content areas (71, 84).

2. Comments by teachers on specific weaknesses and strengths of students have been reported to enhance children's learning. Positive rather than negative comments are most beneficial (82).

3. Work samples which illustrate a child's skills, accompanied by an explanatory note from the teacher, can promote parents' understanding of marks and school objectives, particularly in the early school years (38).

4. Report cards should give enough information to convey the student's status, but they should be functional enough to allow the teacher to mark the student objectively. The more entries appearing on a report card and the greater the range of child behavior covered, the more likely it seems to be that the report card will meet the objectives and definitions of the marking system (20).

5. Dual reporting systems seem to have advantages in providing a more useful picture of the pupil's status and progress (5).

6. Informal letters and parents' conferences enhance the school-parent-child relationship (1: 569).

7. No single system of marking seems to be adequate for reporting; a combination of reporting devices is desirable (24).

In the not inconsiderable research devoted to report cards alone, nearly all of which is largely a matter of studying current practices, one may discover certain trends which may be of value to those planning and developing marking and reporting systems:

1. Report cards are developing certain common characteristics, the most important of which is the provision of space to

record school grades and behavioral achievements, and space for the teacher and parents to add comments.

2. The most common frequency of issuance is six times a year, but the frequency is diminishing, with some schools reducing the number to one or two a semester.

3. Increasingly, schools are discovering the need to augment formal procedures with other devices. While the informal letter as a substitute for the report card has lost favor, it continues to be an important supplement. Parent-teacher conferences are increasingly used to supplement the report card.

4. Parents are sharing to a greater extent in report card revisions. Since the report card represents a communication system between home and school, educators realize that a two-way communication device demands the enlighened understanding of parents.

5. Report cards are becoming increasingly uniform within a school system. During the period of experimentation, there was a tendency to permit individual schools to devise a form of reporting that best suited the purposes of the local unit. Today, many school systems use the same report form for comparable grade levels throughout the system.

SELECTED REFERENCES

1. Ahman, J. Stanley, and Glock, Marvin D. Evaluating Pupil Growth. Boston: Allyn and Bacon, 1963. 588 p.

2. Aiken, Lewis R., Jr. "The Grading Behavior of a College Faculty." Educational and Psychological Measurement 23: 319-22; Summer 1963.

3. Aiken, Lewis R., Jr. "Interdepartmental Variability and Student Expectations of College Grades." Educational and Psychological Measurement 24: 823-29; Winter 1964.

4. Alexander, Eugene D. "The Marking System and Poor Achievement." Teachers College Journal 36: 110-13; December 1964.

5. Alpren, Morton. "A Fair Grading System." Clearing House 35: 113-14; October 1960.

6. Association for Supervision and Curriculum Development. Evaluation as Feedback and Guide. 1967 Yearbook. Washington, D. C.: the Association, 1967. 282 p.

7. Atkinson, John W. An Introduction to Motivation. New York: Van Nostrand, 1964. 335 p.

8. Atkinson, John W., and Feather, Norman T., editors. A Theory of Achievement Motivation. New York: John Wiley and Sons, 1966. 392 p.

9. Austin, Mary C. "Report Cards and Parents." Reading Teacher 18: 660-63; May 1965.

10. Baker, Robert L., and Doyle, Roy. "Teacher Knowledge of Pupil Data and Marking Practices at the Elementary School Level." Personnel and Guidance Journal 37: 644-47; May 1959.

11. Baughman, M. Dale, and Schoonmaker, David. "Grouping Practices in Junior High Schools." Clearing House 36: 111-14; December 1961.

12. Bebell, Clifford F. S. "The Evaluation We Have." Evaluation as Feedback and Guide. 1967 Yearbook. (Edited by Fred T. Wilhelms.) Washington, D. C.: Association for Supervision and Curriculum Development, 1967. Chapter 2, p. 18-46.

13. Boehm, Ann E., and White, Mary Alice. "Pupil's Perceptions of School Marks." Elementary School Journal 67: 237-40; February 1967.

14. Bostrom, Robert N.; Vlandis, John W.; and Rosenbaum, Milton E. "Grades as Reinforcing Contingencies and Attitude Change." Journal of Educational Psychology 52: 112-15; April 1961.

15. Bowers, John E. "A Test of Variation in Grading Standards." Educational and Psychological Measurement 27: 429-30; Summer 1967.

16. Boyd, Margaret. "School Standards in Promotion: Testing, Reporting, Grading." Theory into Practice 4: 95-98; June 1965.

17. Burke, Ronald J. "Some Preliminary Data on the Use of Self-Evaluation and Peer Ratings in Assigning University Course Grades." Journal of Educational Research 62: 444-48; July-August 1969.

18. Burke, Ronald J. "Student Reactions to Course Grades." Journal of Experimental Education 36: 11-13; Summer 1968.

19. Carter, Robert S. "How Invalid Are Marks Assigned by Teachers?" Journal of Educational Psychology 43: 218-19; April 1952.

20. Chansky, Norman M. "Elementary School Teachers Rate Report Card." Journal of Educational Research 56: 523-28; July-August 1963.

21. Chansky, Norman M. "A Note on the Grade Point Average in Research." Educational and Psychological Measurement 24: 95-99; Spring 1964.

22. Chansky, Norman M. "Preferred Items on Pupils' Report Cards." Education 86: 169-73; November 1965.

23. Chansky, Norman M. "Report Card Choices of Gifted Junior High School Pupils." Gifted Child Quarterly 7: 64-65; Summer 1963.

24. Chansky, Norman M. "Report on Parents." NEA Journal 52: 14-15; December 1963.

25. Chansky, Norman M. "X-ray of the School Mark." Educational Forum 26: 347-52; March 1962.

26. Christensen, Donald J. "The Effect of Discontinued Grade Reporting on Pupil Learning." Arithmetic Teacher 15: 724-26; December 1968.

27. Clark, D. Cecil. "Competition for Grades and Graduate-Student Performance." Journal of Educational Research 62: 351-54; April 1969.

28. Crosby, Muriel E. "Good Report." NEA Journal 51: 45-47; April 1962.

29. DePencier, Ida B. "Trends in Reporting Pupil Progress in the Elementary Grades." Elementary School Journal 2: 519-23; May 1951.

30. D'Evelyn, Katherine E. Reporting to Parents. Elementary Instructional Service. Washington, D. C.: Department of Elementary-Kindergarten-Nursery Education, National Education Association, November 1965. 4 p.

31. Doak, E. Dale. "Grading: A Deterrent to Learning." Clearing House 37: 245-48; December 1962.

32. Domino, George. "Differential Prediction of Academic Achievement in Conforming and Independent Settings." Journal of Educational Psychology 59: 256-60; August 1968.

33. Ebel, Robert L. Measuring Educational Achievement. Englewood Cliffs, N. J.: Prentice-Hall, 1965. p. 405.

34. Elton, Charles F. "Three-Year High School Average as a Predictor of College Success." College and University 40: 165-67; Winter 1965.

35. Farwell, Gaylord H.; Nelson, Robert H.; Thompson, Michael L. "Pressures Behind the Grade." Clearing House 38: 462-66; April 1964.

36. Flanders, Ned A.; Morrison, Betty M.; and Brode, E. Leland. "Changes in Pupil Attitudes During the School Year." Journal of Educational Psychology 50: 334-38; October 1968.

37. French, John W. "Comparative Prediction of College Major-Field Grades by Pure-Factor Aptitude, Interest, and Personality Measures." Educational and Psychological Measurement 23: 767-74; Winter 1963.

38. Frenkel, Harriet. "Individualized Report Cards." Instructor 75: 38; September 1965.

39. Garvin, Alfred D. "A Comparison of Students' Predictions of Rank Order on Proximate and Remote Tasks." Journal of Educational Research 61: 176-78; December 1967.

40. Gilbert, June. "What About Our Grading System?" Grade Teacher 74: 18-19; February 1957.

41. Gillcrist, William A., Jr. "A Study of Grading in Arithmetic." Education 86: 177-81; November 1965.

42. Giusti, J. P. "High School Average as a Predictor of College Success: A Survey of the Literature." College and University 39: 200-209; Winter 1964.

43. Gore, Lillian L., and Koury, Rose E. Survey of Early Elementary Education in Public Schools, 1960-61. U. S. Department of Health, Education, and Welfare, Office of Education, Circular No. 782. Washington, D. C.: Government Printing Office, 1965. Chapter 9, "Reporting to Parents," p. 62-65.

44. Halliwell, Joseph W. "Dangers Inherent in Correlating Averages." Journal of Educational Research 55: 327-29; April 1962.

45. Halliwell, Joseph W. "Parental Interpretation of and Reaction to Dual Report Cards." Clearing House 36: 245-47; December 1961.

46. Halliwell, Joseph W. "The Relationship of Certain Factors in Marking Practices in Individualized Reporting Programs." Journal of Educational Research 54: 76-78; October 1960.

47. Halliwell, Joseph W., and Robitaille, Joseph P. "The Relationship Between Theory and Practices in a Dual Reporting Program." Journal of Educational Research 57: 137-41; November 1963.

48. Hanson, Earl D. "Do Boys Get a Square Deal in School?" Education 79: 597-98; May 1959.

49. Hartnett, Rodney T., and Stewart, Clifford T. "Personality Rigidity of Students Showing Consistent Discrepancies Between Instructor Grades and Term-End Examination Grades." Educational and Psychological Measurement 25: 1111-15; Winter 1965.

50. Haskell, John M. "Pass-Fail? A System Worth Trying." Clearing House 42: 172-73; November 1967.

51. Hausdorff, H., and Farr, S. D. "Effect of Grading Practices on the Marks of Gifted Sixth Grade Children." Journal of Educational Research 59: 169-72; December 1965.

52. Hawk, T. L., and DeRidder, L. M. "A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students." Journal of Educational Research 56: 548-50; July-August 1963.

53. Herman, William L. "Teaching Attitude as Related to Academic Grades and Athletic Ability of Prospective Physical Education Teachers." Journal of Educational Research 61: 40-42; September 1967.

54. Hockstad, Patricia. "Report Cards--Helpful or Harmful." Education 84: 174-75; November 1963.

55. Holland, John L. "Some Limitations of Teacher Ratings as Predictors of Creativity." Journal of Educational Psychology 50: 219-23; October 1951.

56. Humphreys, Lloyd G. "The Fleeting Nature of the Prediction of College Academic Success." Journal of Educational Psychology 50: 375-80; October 1968.

57. Jarrett, Calvin D. "Marking and Reporting Practices in the American Secondary School." Peabody Journal of Education 41: 36-48; July 1963.

58. Johnsen, Mauritz, Jr. "Solving the Mess in Marks." New York State Education 49: 12-13, 30; November 1961.

59. Jones, John A. "Grading, Marking, and Reporting in the Modern Elementary School." Educational Forum 19: 45-54; November 1954.

60. Julian, Stanley. Measurement in Today's Schools. Englewood Cliffs, N. J.: Prentice-Hall, 1964. p. 329.

61. Kamii, Constance K., and Weikart, David P. "Marks, Achievement, and Intelligence of Seventh Graders Who Were Retained (Nonpromoted) Once in the Elementary School." Journal of Educational Research 56: 452-59; May-June 1963.

62. Karlins, Marvin; Kaplan, Martin; and Stuart, William. "Academic Attitudes and Performance as a Function of Differential Grading Systems: An Evaluation of Princeton's Pass-Fail System." Journal of Experimental Education 37: 38-50; Spring 1969.

63. Keefer, Karl E. "Self-Prediction of Academic Achievement by College Students." Journal of Educational Research 63: 53-56; October 1969.

64. Kelly, Samuel P., and Thompson, Ralph. "Grading and the Nature of the Discipline." Journal of Higher Education 39: 517-18: December 1968.

65. Kingston, Albert J. "Assigning Grades to Students in Special Reading Classes." Journal of Reading 10: 39-42; October 1966.

66. Kingston, Albert J., and Wash, James A., Jr. "Research on Reporting Systems." National Elementary Principal 45: 36-40; May 1966.

67. Kirby, Bernard. "Three Error Sources in College Grading." Journal of Experimental Education 31: 212-18; December 1962.

68. Kremer, Bruce J. "Is Coeducation Unfair to Boys?" Catholic School Journal 65: 37-39; October 1965.

69. Krutchkoff, Richard G. "The Separation-Level of Grades on a Multiple-Choice Examination." Journal of Experimental Education 36: 63-68; Fall 1967.

70. Lindquist, E. F. "An Evaluation of a Technique for Scaling High School Grades To Improve Prediction of College Success." Educational and Psychological Measurement 23: 623-46; Winter 1963.

71. Ludeman, W. W. "Overhauling School Evaluation." American School Board Journal 140: 37; February 1960.

72. McCowen, Emeline, and Bryan, Roy C. "Reporting to Parents on Pupil Progress." Elementary School Journal 56: 32-34; September 1955.

73. MacDonald, James B., and Clements, H. Millard. "Moral Concerns in Assessing Pupil Growth." National Elementary Principal 45: 29-33: May 1966.

74. Mann, Lester, and others. "Comparison of Formal and Informal Reporting Systems in a First Grade Population." Journal of Educational Research 60: 75-79; October 1966.

75. Mannello, George. "College Teaching Without Grades." Journal of Higher Education 35: 328-34; June 1964.

76. Miner, Betty Crowether. "Three Factors of School Achievement." Journal of Educational Research 60: 370-76; April 1967.

77. Morris, Lucille. "Evaluating and Reporting Pupil Progress." Elementary School Journal 53: 144-49; November 1952.

78. Murstein, Bernard I. "The Relationship of Grade Expectations and Grades Believed To Be Deserved to Actual Grades Received." Journal of Experimental Education 33: 357-62; Summer 1965.

79. National Education Assocciation, Research Division. Kindergarten Education in Public Schools, 1967-68. Research Report 1969-R6. Washington, D. C.: the Association, 1969. p. 31.

80. National Education Association, Research Division. "Reporting Pupil Progress." NEA Research Bulletin 47: 75-76; October 1969.

81. Odell, Charles W. Educational Measurement in the High School. New York: Century Co., 1930. 641 p.

82. Page, Ellis B. "Teacher Comments and Student Performances: A Seventy-four Classroom Experiment in School Motivation." Journal of Educational Psychology 49: 173-81; August 1958.

83. Pemberton, John H. "Rx for Report Card Blues." Clearing House 36: 75-77; October 1961.

84. Phelps, H. V. "How To Design a Report Card Parents Can Understand." School Management 8: 72-74; May 1964.

85. Philbrick, Joseph L., and O'Donnell, Patrick I. "Precision in Grading Practices--Panacea or Problem?" Journal of Educational Research 62: 173-76; December 1968.

86. Pigge, Fred L. "Teaching Effectiveness of 'A' and 'C' Elementary Teachers." Journal of Educational Research 62: 99-102; November 1968.

87. Quinn, George R., and Szuberla, Charles A. "Relative Grading." Clearing House 37: 490-94; April 1963.

88. Reavis, William C. "Report Cards." School Review 60: 199-200; April 1952.

89. Richards, James M., Jr., and Taylor, Calvin W. "Predicting Academic Achievement in a College of Medicine from Grades, Test Scores, Interviews, and Ratings." Educational and Psychological Measurement 21: 987-94; Winter 1961.

90. Richardson, Sybil. "How Do Children Feel About Reports to Parents?" California Journal of Elementary Education 24: 98-111; November 1955.

91. Richardson, Sybil. "Reporting to Parents." Instructor 69: 9; June 1960.

92. Rinsland, Henry D. Constructing Tests and Grading in Elementary and High School. New York: Prentice-Hall, 1937. 323 p.

93. Robertson, Malcolm H. "Counselor and Student Estimates of Grades as Predictors of Academic Achievement." Journal of Educational Research 54: 73-75; October 1960.

94. Rolf, Fred J. "Reporting to Parents." Ohio Schools 39: 10-11; May 1961.

95. Rothney, John W. M. Evaluating and Reporting Pupil Progress. What Research Says to the Teacher, No. 7. Washington, D. C.: National Education Association, 1955. 33 p.

96. Russell, David H. "What Does Research Say About Self-Evaluation?" Journal of Educational Research 46: 561-73; April 1953.

97. Russell, I. L., and Talman, W. H. "Personality: Does It Influence Teachers' Marks?" Journal of Educational Research 48: 561-64; April 1965.

98. Scales, Eldridge E. "Variability in Grading Practices of Instructors of a Multiple-Section Natural Science Course." Science Education 50: 332-35; October 1966.

99. Schwartz, Alfred, and Tiedeman, Stuart. Evaluating Student Progress in the Secondary School. New York: Longmans, Green and Co., 1957. 434 p.

100. Sgan, Mathew R. "The First Year of Pass-Fail at Brandeis University: A Report." Journal of Higher Education 40: 135-44; February 1969.

101. Smith, Ann Z., and Dobben, John E. "Marks and Marking Systems." Encyclopedia of Educational Research. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. p. 783-91.

102. Smith, Eugene R., and Tyler, Ralph. Appraising and Recording Student Progress. Progressive Education Association, Adventure in American Education, Volume 3. New York: Harper and Brothers, 1942. 550 p.

103. Stallings, William M.; Wolff, Joseph L.; and Maehr, Martin L. "Fear of Failure and the Pass-Fail Grading Option." Journal of Experimental Education 38: 87-91; Winter 1969.

104. Storey, Arthur G. "Some Evidence in the Essay-Item Case." Journal of Educational Research 61: 351-54; April 1968.

105. Storey, Arthur G. "The Versatile Multiple-Choice Item." Journal of Educational Research 62: 169-72; December 1968.

106. Strang, Ruth. How To Report Pupil Progress. Chicago: Science Research Associates, 1955. 47 p.

107. Stricker, George, and Huber, J. T. "The Graduate Record Examination and Undergraduate Grades as Predictors of Success in Graduate School." Journal of Educational Research 60: 466-68; July-August 1967.

108. Stroup, Francis. "The Grade Point Average Is Obsolete." Journal of Higher Education 34: 10-15; January 1963.

109. Stroup, Francis. "Grouping Errors in the Grade-Point Average." Journal of Experimental Education 34: 31-33; Spring 1966.

110. Suehr, John J. "Marking and Reporting Practice in a Modern High School." Bulletin of the National Association of Secondary-School Principals 46: 29-32; April 1962.

111. Symonds, Percival M. Measurement in Secondary Education. New York: Macmillan Co., 1928. 588 p.

112. Symonds, Percival M. "Pupil Evaluation and Self-Evaluation." Teachers College Record 54: 138-49; December 1952.

113. Taylor, Bob L., and McKean, Robert C. "Divergent Thinkers and Teacher Education." Journal of Educational Research 61: 417-18; May-June 1968.

114. Terwilliger, James S. "Self-Reported Marking Practices and Policies in Public Secondary Schools." Bulletin of the National Association of Secondary-School Principals 50: 5-37; March 1966.

115. Thomas, R. Murray. Judging Student Progress. New York: Longmans, Green and Co., 1960. 472 p.

116. Vredroe, Lawrence C., and Lindecamp, Charles D. "How Shall We Make the Recording and Reporting of Pupil Progress More Meaningful?" Bulletin of the National Association of Secondary-School Principals 37: 179-85; April 1953.

117. Wagner, G. W. "What Schools Are Doing: Reporting Pupil Progress." Education 80: 381; February 1960.

118. Wall, W. A., and Miller, K. M. "Motivation and Counter-Motivation." Proceedings of the XIV International Congress of Applied Psychology 3: 161-75; 1962.

119. Wexley, Kenneth N.; McLaughlin, Gerald W.; and Montgomery, James R. "Relationships Between Different Variables and Performances in a Junior-Year English Examination." Journal of Experimental Education 36: 58-62; Fall 1967.

120. Willingham, Warren W. "The Effect of Grading Variations upon the Accuracy of Predicting Freshman Grades." College and University 40: 159-64; Winter 1965.

121. Wrinkle, William L. Improving Marking and Reporting Practices in Elementary and Secondary Schools. New York: Holt, Rinehart and Winston, 1947. 120 p.

122. Yauch, Wilbur A. "School Marks and Their Reporting." NEA Journal 5: 50-58; May 1961.

Research Reports

1968-R3 Head Start Programs Operated by Public School Systems, 1966-67. 42 p. $1.00. #435-13346.

1968-R5 Salary Schedules for Principals, 1967-68. 126 p. $2.50. #435-13350.

1968-R6 Nursery School Education, 1966-67. 48 p. $1.00. #435-13352.

1968-R7 Salaries in Higher Education, 1967-68. 92 p. $1.50. #435-13354.

1968-R11 Selected Statistics of Local School Systems, 1966-67. 119 p. $2.25. #435-13362.

1968-R15 State Minimum Salary Laws for Teachers, 1968-69. 43 p. $1.00. #435-13372.

1968-R17 Faculty Salary Schedules in Community-Junior Colleges, 1967-68. 50 p. $1.00. #435-13376.

1969-R2 Faculty Salary Schedules in Colleges and Universities, 1967-68. 48 p. $1.00. #435-13384.

1969-R6 Kindergarten Education in Public Schools, 1967-68. 56 p. $1.25. #435-13394.

1969-R7 24th Biennial Salary Survey of Public-School Professional Personnel, 1969. 346 p. $5.00. #435-13396.

1969-R8 Grievance Procedures for Teachers in Negotiation Agreements. 37 p. $1.00. #435-22882.

1969-R9 Paid Leave Provisions for Teachers in Negotiation Agreements. 95 p. $1.75. #435-22884.

1969-R10 The Pupil's Day in Court: Review of 1968. 65 p. $1.25. #435-22886.

1969-R11 The Teacher's Day in Court: Review of 1968. 61 p. $1.25. #435-22888.

1969-R12 High Spots in State School Legislation, January 1-August 31, 1969. 122 p. $2.25. #435-22890.

1969-R13 Salary Schedules for Teachers, 1969-70. 122 p. $2.25. #435-22892.

1969-R14 Teacher Supply and Demand in Public Schools, 1969. 77 p. $1.50. #435-22894.

1969-R15 Estimates of School Statistics, 1969-70. 38 p. $1.00. #435-22896.

1970-R1 Rankings of the States, 1970. 74 p. $1.50. #435-25430.

1970-R2 Maximum Salaries Scheduled for School Administrators, 1969-70. 98 p. $2.00. #435-25432.

1970-R3 Economic Status of the Teaching Profession, 1969-70. 68 p. $1.25. #435-25434.

1970-R4 Salary Schedule Supplements for Extra Duties, 1969-70. 66 p. $1.25. #435-25436.

Research Summaries

1966-S1 Inservice Education of Teachers. 19 p. 60¢. #434-22802.

1966-S2 Homework. 12 p. 30¢. #434-22804.

1967-S1 School Dropouts. 55 p. 75¢. #434-22808.

1968-S1 Class Size. 49 p. $1.00. #434-22810.

1968-S2 The Rescheduled School Year. 38 p. $1.00. #434-22812.

1968-S3 Ability Grouping. 52 p. $1.00. #434-22814.

1970-S1 Marking and Reporting Pupil Progress. 55 p. $1.25. #434-22854.

Research Memos

1968-16 Adult Education Statistics, 1968-69. December 1968. 18 p. 50¢. #431-22842.

1969-8 Index Salary Schedules for Teachers, 1968-69. March 1969. 15 p. 45¢. #431-22822.

1969-10 Professional Growth Requirements Specified in 1968-69 Salary Schedules. May 1969. 15 p. 45¢. #431-22832.

1969-11 Use of Teacher Aides, 1968-69. May 1969. 15 p. 45¢. #431-22830.

1969-12 Merit Provisions in Teachers' Salary Schedules, 1968-69. April 1969. 11 p. 35¢. #431-22826.

1969-13 Salary Schedule Provisions for Non-Clerical Support Personnel, 1968-69. July 1969. 21 p. 60¢. #431-22856.

1969-19 Salary Schedule Provisions for Full-Time Guidance Counselors, 1968-69. July 1969. 33 p. 75¢. #431-22868.

1969-20 Scheduled Salaries for School Office Personnel, 1968-69, Reporting Systems with Enrollments of 6,000 or More. July 1969. 17 p. 50¢. #431-22870.

1969-21 Index to NEA Research Division Publications Currently Available. August 1969. 22 p. 60¢. #431-22872.

1969-22 Leaves of Absence Provisions for Teachers, 1968-69. October 1969. 23 p. 65¢. #431-22874.

1969-25 Salaries Paid Secretarial and Clerical Personnel in Public Schools, 1966-67 Through 1968-69. October 1969. 11 p. 35¢. #431-22880.

1969-26 Minimum Annual Salaries for Teachers. December 1969. 3 p. 15¢. #431-22900.

1969-27 Teacher Strikes and Work Stoppages. December 1969. 11 p. 35¢. #431-22902.

1970-1 Average Salary of Instructional Staff in Public Schools, by State: 1939-40 to 1969-70. January 1970. 1 p. 10¢. #431-22906.

1970-3 Beginning Salaries for College Graduates, June 1970. January 1970. 7 p. 25¢. #431-22904.

1970-4 Placement of Newly Appointed Experienced Teachers on Salary Schedules, 1969-70. April 1970. 38 p. 80¢. #431-22910.

1970-6 Pupil-Staff Ratios, 1968-69. March 1970. 14 p. 40¢. #431-22908.

1970-7 Merit Provisions in Teachers' Salary Schedules, 1969-70. April 1970. 9 p. #431-22912.